

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2005, 09:51:36 ; Search time 39.4398 Seconds  
(without alignments)  
1118.605 Million cell updates/sec

Title: US-10-725-121A-103

Perfect score: 3090  
Sequence: 1 MFGRKKRVEISAPSNFHEH.....LAKAGPPASIVLMRQNRTR 591

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgm2\_6/prodata/1/1aa/5A\_COMB.pep.\*  
2: /cgm2\_6/prodata/1/1aa/5B\_COMB.pep.\*  
3: /cgm2\_6/prodata/1/1aa/6A\_COMB.pep.\*  
4: /cgm2\_6/prodata/1/1aa/6B\_COMB.pep.\*  
5: /cgm2\_6/prodata/1/1aa/PCTUS\_COMB.pep.\*  
6: /cgm2\_6/prodata/1/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3090	100.0	591	3	US-09-082-737-2
2	3090	100.0	591	4	US-09-688-188B-103
3	3090	100.0	591	4	US-09-718-032-2
4	3090	100.0	591	4	US-09-291-417D-103
5	3090	100.0	591	4	US-09-949-016-6665
6	3090	100.0	620	4	US-09-949-016-7206
7	2073	67.1	398	4	US-09-688-188B-30
8	2073	67.1	398	4	US-09-291-417D-30
9	1473.5	47.7	681	4	US-09-688-188B-29
10	1473.5	47.7	681	4	US-09-291-417D-29
11	1473.5	47.7	694	4	US-09-949-016-10445
12	1470.5	47.6	681	4	US-09-765-815-2
13	1277	41.3	250	3	US-09-082-737-3
14	1277	41.3	250	4	US-09-718-032-3
15	912	29.5	544	2	US-08-935-760-2
16	909	29.4	544	4	US-09-949-016-11562
17	907	29.4	544	4	US-09-688-188B-95
18	907	29.4	544	4	US-09-291-417D-95
19	904	29.3	524	4	US-09-538-092-1301
20	900	29.1	524	4	US-08-615-942A-2
21	900	29.1	524	4	US-09-237-325-2
22	899	29.1	544	3	US-08-559-397A-19
23	894	28.9	506	1	US-08-369-780-2
24	894	28.9	506	1	US-08-475-682-2
25	894	28.9	506	1	US-08-780-833-2
26	894	28.9	506	1	US-08-636-036-2
27	894	28.9	506	3	US-08-918-509-2

28	894	28.9	506	3	US-09-108-262-2	Sequence 2, Appl
29	894	28.9	506	4	US-09-688-188B-94	Sequence 94, Appl
30	894	28.9	506	4	US-09-291-417D-94	Sequence 94, Appl
31	893.5	28.9	551	4	US-09-949-016-10951	Sequence 10951, A
32	892.5	28.9	545	4	US-09-538-092-1297	Sequence 1297, Ap
33	891.5	28.9	545	2	US-08-935-760-4	Sequence 4, Appl
34	891.5	28.9	545	4	US-09-688-188B-93	Sequence 93, Appl
35	891.5	28.9	545	4	US-09-291-417D-93	Sequence 93, Appl
36	867	28.1	544	3	US-08-559-397A-30	Sequence 29, Appl
37	849	27.5	544	3	US-08-559-397A-30	Sequence 30, Appl
38	833.5	27.0	465	2	US-08-114-555A-2	Sequence 2, Appl
39	821	26.6	465	3	US-08-559-397A-31	Sequence 31, Appl
40	814.5	26.4	465	3	US-08-559-397A-2	Sequence 2, Appl
41	804	26.0	305	4	US-09-765-815-10	Sequence 10, Appl
42	782	25.3	793	3	US-09-588-256-10	Sequence 3, Appl
43	769	24.9	268	2	US-08-852-743-3	Sequence 3, Appl
44	769	24.9	268	3	US-09-185-370-3	Sequence 3, Appl
45	755	24.4	410	4	US-09-248-796A-18461	Sequence 18461, A

## ALIGNMENTS

```

RESULT 1
US-09-082-737-2
; Sequence 2, Application US/09082737
; Patent No. 6013500
;
GENERAL INFORMATION:
; APPLICANT: Minden, Audrey
; TITLE OF INVENTION: PAK1 A No. 6013500e1 Gene Encoding A Serine/
; TITLE OF INVENTION: Threonine Kinase
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11230
;
COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/082,737
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/55311
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 591 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-09-082-737-2
Query Match 100.0%; Score 3090; DB 3; Length 591;
Best Local Similarity 100.0%; Pred. No. 3e-191;
Matches 591; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MFGRKKRVEISAPSNFHEHVTGPDHOKFTGLPRQWOSTLIBSARRKPLVDPACIT 60
Db 1 MFGRKKRVEISAPSNFHEHVTGPDHOKFTGLPRQWOSTLIBSARRKPLVDPACIT 60
QY 61 SIQPAKPTTVKSGAKGGLTLIDDFRNNSVTSNSLRKRSPPPPARAQDNGMPEP 120

```

THIS PAGE BLANK (USPTO)

```
Db 61 S10PGAPKTIVRGSGKAKDGA1TL1LDBFENMSVTRNSILRDSPPPPARARQENGMPEE 120
Qy 121 PATTARAGPGKAGSRGRFAGHSBAGGSGDRBRAGPEKRPKSSRBSGSGPQESSRDRLPL 180
Db 121 PATTARAGPGKAGSRGRFAGHSBAGGSGDRBRAGPEKRPKSSRBSGSGPQESSRDRLPL 180
Qy 181 SGPDVGTPOGAGLASGAKLAAGRPFNTYPRADTDHPSRGAQGBPHDVAENGPASAGLAIP 240
Db 181 SGPDVGTPOGAGLASGAKLAAGRPFNTYPRADTDHPSRGAQGBPHDVAENGPASAGLAIP 240
Qy 241 QSSSSSRPPTRRAGAPSPGVLPFHASBPOLAPACTPAPAAVGPFGPSPQREPORVS 300
Db 241 QSSSSSRPPTRRAGAPSPGVLPFHASBPOLAPACTPAPAAVGPFGPSPQREPORVS 300
Qy 301 HEQFRALQVVDPGDRSYLDNFIKIGSGSTGIVCIATVRSSGKLVAVKMDLRKQORR 360
Db 301 HEQFRALQVVDPGDRSYLDNFIKIGSGSTGIVCIATVRSSGKLVAVKMDLRKQORR 360
Qy 361 ELLENFVIMRDYOHENNVEMNYSYLVGDELMVMEELLEGALTDIYTHTRMNEBOJIAAV 420
Db 361 ELLENFVIMRDYOHENNVEMNYSYLVGDELMVMEELLEGALTDIYTHTRMNEBOJIAAV 420
Qy 421 CLAVLOALSVLAHQVTHRDIKSDSILLTHDGRVKSDFGFCQVSKVPRKSLVGTPT 480
Db 421 CLAVLOALSVLAHQVTHRDIKSDSILLTHDGRVKSDFGFCQVSKVPRKSLVGTPT 480
Qy 481 WMAPELISRLPYGPBYDIMSIGIMVTEMVDGEPYFNEBPPLKAMKIMRDLPPRLKNYLHK 540
Db 481 WMAPELISRLPYGPBYDIMSIGIMVTEMVDGEPYFNEBPPLKAMKIMRDLPPRLKNYLHK 540
Qy 541 VSPSLGFLDRLVLRDPAQRATAELKHPFLAKAGPPASIVPLMRQNRTR 591
Db 541 VSPSLGFLDRLVLRDPAQRATAELKHPFLAKAGPPASIVPLMRQNRTR 591

RESULT 2
US-09-688-188B-103
; Sequence 103, Application US/09688188B
; Patent No. 6656716
; GENERAL INFORMATION:
; APPLICANT: PIOMMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; FILE REFERENCE: 038602/0328
; CURRENT APPLICATION NUMBER: US/09/688,188B
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 09/291,417
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 60/081,784
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 103
; LENGTH: 591
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-688-188B-103

Query Match 100.0%; Score 3090; DB 4; Length 591;
Best Local Similarity 100.0%; Pred. No. 3e-191;
Matches 591; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 MFGKRRKRVIEISAPSNFHRVHTGPDQHEOKFTGLPRQOSILBESARRPKPLVDPACIT 60
Db 1 MFGKRRKRVIEISAPSNFHRVHTGPDQHEOKFTGLPRQOSILBESARRPKPLVDPACIT 60
Qy 61 S10PGAPKTIVRGSGKAKDGA1TL1LDBFENMSVTRNSILRDSPPPPARARQENGMPEE 120
Db 61 S10PGAPKTIVRGSGKAKDGA1TL1LDBFENMSVTRNSILRDSPPPPARARQENGMPEE 120
Qy 121 PATTARAGPGKAGSRGRFAGHSBAGGSGDRBRAGPEKRPKSSRBSGSGPQESSRDRLPL 180
Db 121 PATTARAGPGKAGSRGRFAGHSBAGGSGDRBRAGPEKRPKSSRBSGSGPQESSRDRLPL 180
```

```
Db 121 PATTARAGPGKAGSRGRFAGHSBAGGSGDRBRAGPEKRPKSSRBSGSGPQESSRDRLPL 180
Qy 181 SGPDVGTPOGAGLASGAKLAAGRPFNTYPRADTDHPSRGAQGBPHDVAENGPASAGLAIP 240
Db 181 SGPDVGTPOGAGLASGAKLAAGRPFNTYPRADTDHPSRGAQGBPHDVAENGPASAGLAIP 240
Qy 241 QSSSSSRPPTRRAGAPSPGVLPFHASBPOLAPACTPAPAAVGPFGPSPQREPORVS 300
Db 241 QSSSSSRPPTRRAGAPSPGVLPFHASBPOLAPACTPAPAAVGPFGPSPQREPORVS 300
Qy 301 HEQFRALQVVDPGDRSYLDNFIKIGSGSTGIVCIATVRSSGKLVAVKMDLRKQORR 360
Db 301 HEQFRALQVVDPGDRSYLDNFIKIGSGSTGIVCIATVRSSGKLVAVKMDLRKQORR 360
Qy 361 ELLENFVIMRDYOHENNVEMNYSYLVGDELMVMEELLEGALTDIYTHTRMNEBOJIAAV 420
Db 361 ELLENFVIMRDYOHENNVEMNYSYLVGDELMVMEELLEGALTDIYTHTRMNEBOJIAAV 420
Qy 421 CLAVLOALSVLAHQVTHRDIKSDSILLTHDGRVKSDFGFCQVSKVPRKSLVGTPT 480
Db 421 CLAVLOALSVLAHQVTHRDIKSDSILLTHDGRVKSDFGFCQVSKVPRKSLVGTPT 480
Qy 481 WMAPELISRLPYGPBYDIMSIGIMVTEMVDGEPYFNEBPPLKAMKIMRDLPPRLKNYLHK 540
Db 481 WMAPELISRLPYGPBYDIMSIGIMVTEMVDGEPYFNEBPPLKAMKIMRDLPPRLKNYLHK 540
Qy 541 VSPSLGFLDRLVLRDPAQRATAELKHPFLAKAGPPASIVPLMRQNRTR 591
Db 541 VSPSLGFLDRLVLRDPAQRATAELKHPFLAKAGPPASIVPLMRQNRTR 591

RESULT 3
US-09-718-032-2
; Sequence 2, Application US/09718032
; Patent No. 6667168
; GENERAL INFORMATION:
; APPLICANT: The Trustees of Columbia University
; APPLICANT: Manden, Audrey
; TITLE OF INVENTION: PAK1, A NOVEL GENE ENCODING A SERINE/THREONINE KINASE
; FILE REFERENCE: 575/55311-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/718,032
; PRIOR FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: PCT/US99/11341
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 09/082,737
; PRIOR FILING DATE: 1998-05-21
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 591
; TYPE: PRT
; ORGANISM: human
US-09-718-032-2

Query Match 100.0%; Score 3090; DB 4; Length 591;
Best Local Similarity 100.0%; Pred. No. 3e-191;
Matches 591; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 MFGKRRKRVIEISAPSNFHRVHTGPDQHEOKFTGLPRQOSILBESARRPKPLVDPACIT 60
Db 1 MFGKRRKRVIEISAPSNFHRVHTGPDQHEOKFTGLPRQOSILBESARRPKPLVDPACIT 60
Qy 61 S10PGAPKTIVRGSGKAKDGA1TL1LDBFENMSVTRNSILRDSPPPPARARQENGMPEE 120
Db 61 S10PGAPKTIVRGSGKAKDGA1TL1LDBFENMSVTRNSILRDSPPPPARARQENGMPEE 120
Qy 121 PATTARAGPGKAGSRGRFAGHSBAGGSGDRBRAGPEKRPKSSRBSGSGPQESSRDRLPL 180
Db 121 PATTARAGPGKAGSRGRFAGHSBAGGSGDRBRAGPEKRPKSSRBSGSGPQESSRDRLPL 180
Qy 181 SGPDVGTPOGAGLASGAKLAAGRPFNTYPRADTDHPSRGAQGBPHDVAENGPASAGLAIP 240
Db 181 SGPDVGTPOGAGLASGAKLAAGRPFNTYPRADTDHPSRGAQGBPHDVAENGPASAGLAIP 240
```



**THIS PAGE BLANK (uspto)**

## Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

**When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2005, 10:09:48 ; Search time 92.1557 Seconds  
(without alignments)  
2211.556 Million cell updates/sec

Title: US-10-725-121A-30

Sequence: 1 ASGAKLAAGPFTYPRADT.....LAKAGPPASIVPLKQRNTR 398

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues  
Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot 03:.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2073	100.0	591	1	PAK4_HUMAN
2	1963	94.7	407	1	Q96013 mus sapien
3	1963	94.7	593	1	Q8K0T2 PAK4_MOUSE
4	1963	94.7	593	2	Q8K0T2 PAK4_MOUSE
5	1963	94.2	593	2	Q8K0T2 PAK4_MOUSE
6	1926	92.9	501	2	Q8N4E1
7	1925	92.9	438	2	Q8N4E1
8	1925	92.9	467	2	Q9U5S8
9	1612.5	77.8	650	2	Q9U5S8
10	1586	76.5	663	2	Q6GMI1
11	1351	65.2	719	2	Q8C0I5
12	1343	64.8	719	2	Q8BVB0
13	1340	64.6	719	1	PAK7_HUMAN
14	1340	64.6	719	2	Q8T8S3
15	1339.5	64.6	711	2	Q6NV04
16	1333	64.3	719	2	Q6RWS7
17	1259.5	60.8	639	1	PAK1_DROME
18	1241	59.9	540	2	Q8NCS5
19	1237.5	59.7	558	2	Q7QGS5
20	1197	57.7	681	1	PAK6_HUMAN
21	1160	56.0	229	2	Q9CS71
22	911.5	44.0	540	2	Q9U3M1
23	911.5	44.0	542	2	Q18637
24	856	41.3	704	2	Q24190
25	856	41.3	704	2	Q9V1I3
26	851	41.1	704	2	Q24213
27	843	40.7	559	1	PAK3_HUMAN
28	843	40.7	559	1	PAK3_PANTR
29	843	40.7	559	1	PAK3_PONPY
30	842	40.6	524	1	PAK2_HUMAN
31	838	40.4	704	2	Q7QD76

32	833.5	40.2	524	2	Q61SC3	Q61SC3 homo sapien
33	832.5	40.2	544	1	PAK3_RAT	Q62828 rattus norv
34	832.5	40.2	544	1	Q8K1R6	Q8K1R6 mus musculu
35	832.5	40.2	559	1	PAK3_MOUSE	Q61036 mus musculu
36	829.5	40.0	524	1	PAK2_MOUSE	Q8C141 mus musculu
37	829.5	40.0	524	1	PAK2_RAT	Q64303 rattus norv
38	828.5	40.0	524	2	Q9QY10	Q9QY10 rattus norv
39	827.5	39.9	564	2	Q8AXB4	Q8AXB4 xenopus lae
40	825	39.8	524	1	PAK2_RABIT	Q29502 oryctolagus
41	822	39.7	517	2	Q6DG42	Q6DG42 brachydantio
42	822	39.7	517	2	Q8A677	Q8A677 brachydantio
43	817.5	39.4	545	1	PAK1_HUMAN	Q13153 homo sapien
44	817	39.4	544	1	PAK1_RAT	P35465 rattus norv
45	817	39.4	577	2	Q803Z0	Q803Z0 brachydantio

## ALIGNMENTS

RESULT 1	PAK4_HUMAN	STANDARD;	PRT;	591 AA.
ID	PAK4_HUMAN			
AC	Q96013; Q9BUJ3;			
DT	16-OCT-2001 (Ref. 40, Created)			
DT	16-OCT-2001 (Ref. 40, Last sequence update)			
DT	05-JUL-2004 (Ref. 44, Last annotation update)			
DE	Serine/threonine-protein kinase PAK 4 (EC 2.7.1.37) (p21-activated kinase 4) (PAK-4).			
GN	Name=PAK4;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Butcheria; Primates; Carnivora; Hominoidea; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RX	MEDLINE=99043860; PubMed=9822598; DOI=10.1093/emboj/17.22.6527;			
RA	Abo A., Qu J., Camarero M.S., Dan C., Fritsch A., Baud V.,			
RT	Belisle B., Minden A.,			
RT	"PAK4", a novel effector for Cdc42Hs, is implicated in the			
RT	reorganization of the actin cytoskeleton and in the formation of			
RT	filopodia."			
RL	EMBO J. 17:6527-6540(1998).			
RN	[2]			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RA	Meinick M.B.;			
RT	Submitted (May-1997) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).			
RC	TISSUE=Bye, Pancreas, and Placenta;			
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Udell T.B., Toshiyuki S., Cavatini P., Prange C.,			
RA	Rosa S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RA	Bohak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Butterfield Y.S.N., Krzywinski M.I., Skalka U., Smalley D.B.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Falley J., Helton B., Kettman M., Madan A., Rodriguez S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,			
RA	Butterfield Y.S.N., Krzywinski M.I., Skalka U., Smalley D.B.,			
RA	Schneerch A., Schein J.E., Jones S.J.M., Marra M.A.,			
RT	"Generation and Initial Analysis of more than 15,000 full-length human			
RT	and mouse cDNA sequences."			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
CC	-I- FUNCTION: Activates the JNK pathway. Implicated in the			
CC	reorganization of the actin cytoskeleton and in the formation of			
CC	filopodia.			





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: November 17, 2005, 10:00:02 ; Search time 21.3286 Seconds

(without alignments)  
1795.441 Million cell updates/sec

Title: US-10-725-121A-30

Sequence: 1 ASGAKLAAAGPFTYTPRADT.....LAKAGPPASIVPLMRQNR 398

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	911.5	44.0	540	T19956	hypothetical prote
2	911.5	44.0	542	T19952	hypothetical prote
3	833.5	40.2	525	S58682	protein kinase, p2
4	832.5	40.2	544	A57597	beta-p21-activated
5	817.5	39.4	545	G01773	p21-activated prot
6	816	39.4	544	S40482	serine/threonine-s
7	815.5	39.3	544	T49376	p21 activated kina
8	792	38.2	939	S28394	probable serine/th
9	763.5	36.8	658	T39500	serine/threonine-s
10	759	36.6	1230	T18256	probable serine/th
11	759	36.6	1230	T18259	serine/threonine p
12	754.5	36.4	658	S60170	protein kinase pak
13	742.5	35.8	622	T15467	hypothetical prote
14	722	34.8	378	T26684	hypothetical prote
15	672.5	32.4	842	S60402	protein kinase CLA
16	670	32.3	589	T38086	serine/threonine-p
17	607.5	29.3	655	S51884	probable protein k
18	498	24.0	471	T39232	probable serine th
19	494	23.8	1102	TJC6316	probable protein k
20	484	23.3	836	B96716	probable serine/th
21	478	23.0	653	T34356	hypothetical prote
22	476	23.0	819	A53714	protein kinase (BC
23	470.5	22.7	839	T29372	hypothetical prote
24	464	22.4	426	S71886	Ste20-like protein
25	463	22.3	1231	T18532	serine/threonine pr
26	462	22.2	1233	T14157	protein kinase SK2
27	460	22.2	1206	T34021	protein kinase P
28	458.5	22.1	690	C96572	protein P12M16.4 l
29	455.5	22.0	1080	S48944	hypothetical prote

30	455	21.9	1233	2	T30989	serine/threonine p
31	450.5	21.7	312	2	T38525	serine/threonine p
32	445	21.5	1001	2	T17365	serine/threonine p
33	440.5	21.2	1075	2	T27623	hypothetical prote
34	440.5	21.2	1080	2	T27622	hypothetical prote
35	437.5	21.1	1062	2	S46367	protein kinase CDC
36	428	20.6	1228	2	T18697	hypothetical prote
37	421	20.3	553	2	T01479	hypothetical prote
38	413	19.9	561	2	T51417	protein kinase-lik
39	410	19.8	659	1	A39723	protein kinase byr
40	410	19.8	883	2	A96662	hypothetical prote
41	409.5	19.8	652	2	T39722	serine/threonine p
42	408.5	19.7	891	2	T40503	protein kinase kin
43	404.5	19.5	1135	1	A29813	132k ninaC protein
44	404.5	19.5	1501	1	B29813	174k ninaC protein
45	399.5	19.3	891	2	A38903	protein kinase 1 -

## ALIGNMENTS

RESULT 1									
T19956 hypothetical protein C45B11.1b - Caenorhabditis elegans									
C:Species: Caenorhabditis elegans									
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 16-Aug-2004									
C:Accession: T19956									
R:McMurray, A.									
submitted to the EMBL Data Library, June 1996									
A:Reference number: Z19202									
A:Accession: T19956									
A:Status: preliminary; translated from GB/EMBL/DBJ									
A:Molecule type: DNA									
A:Residues: 1-540 <WIL>									
A:Cross-references: UNIPROT:Q9U3M1; EMBL:Z74029; PIDN:CAA98433.1; GSPDB:GN00023; CESP:C4									
A:Experimental source: clone C45B11									
C:Genetic81									
A:Gene: CESP:C45B11.1b									
A:Map position: 5									
A:Intons: 13/2; 62/3; 104/2; 231/3; 367/2; 468/3									
C:Superfamily: protein kinase homology									
Query Match									
Best Local Similarity 49.6%; Pred. No. 1, de-29;									
Matches 202; Conservative 48; Mismatches 104; Indels 53; Gaps 6;									
Score 911.5; DB 2; Length 540;									
QY	31	PHDVAPNGPAGGLAIPSSSSSR-----PPTRARG-----AASP-----	66						
DB	119	PSVARSNSLRISATASVNVSSARHSFRPTLPVSGRGYPRNDPSAFLPLRNQKPM	178						
QY	67	---GVLAGPH-----ASBPQLAPACTPAAPAVGPPGPRSPQRP---ORVS	107						
DB	179	TTTGVEKEPHQYQOITIVAPSRITTPQLQPKS-----PSTPAMKQGPCTGVS	226						
QY	108	HEQFRALQVNDPGDPRSTLDNFKIGESGTGVCIAITRRSGKLVAVKMDLRKQOR	167						
DB	229	DEEFRRALKFVVVVGTPDRSDLTIDYKQIGESTGVEAAVYIKIQIYAAVVRKMLRKQOR	288						
QY	168	ELLPNEVIMRDYQHEHNVEMVNSYLVGDELWVMEBLLEGALTDIYTHRRANBEQIAAV	227						
DB	289	ELLPNEVSIIRKQYQHPHIVAFSSSHLYVDLWVMEBMEGSLTDIYTAIRMTBPQIATI	348						
QY	228	CLAVLQALSVLAQGVIRHDIKSDIILLTHDGVKLSDFGCAQVNSKEVRRKSLVGPY	287						
DB	349	SRQVLGALDPLHARKVIRHDIKSDIILLKRDGVTKLTDFFGCOQLSEVPRRSLSVGPY	408						
QY	288	WMAPELISRLPYGPVDVWSLGMVTEMVDGEPYPRNEPLKAMKMRDLPRLLKTLHK	347						
DB	409	WTAAEVYIAREPYDTRDIDWSFGIMLIEMVGEPPFPNDQFQAMKTRIDREARFSRHAK	468						
QY	348	VSSSLKGFDLRLVVRDPAQATAAELKHPFLAKAGPPASIVPLMRQ 394							
DB	469	VSVLELSLSHCIVQVKNKMPAKDLIRHFPKAKQHSSTIAILLQ 515							

## RESULT 2

T19952

hypothetical protein C45B11.1a - *Caenorhabditis elegans*C/Species: *Caenorhabditis elegans*

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 16-Aug-2004

C/Accession: T19952

R/McMurray, A.

Submitted to the EMBL Data Library, June 1996

A/Reference number: Z19202

A/Accession: T19952

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1542 &lt;M1&gt;

A/Cross-references: UNIPROT:Q18637; EMBL:Z74029; FIDN:CAA98429.1; GSPDB:GN00023; CESP:C4

A/Experimental source: clone C45B11

C/Genetic:

A/Map position: 5

A/Insertions: 13/2; 62/3; 104/2; 233/3; 369/2; 470/3

C/Superfamily: protein kinase homology

Query Match 44.0%; Score 911.5; DB 2; Length 542;  
 Best Local Similarity 49.6%; Pred. No. 1.8e-29;  
 Matches 202; Conservative 46; Mismatches 104; Indels 53; Gaps 6;

31 PHDVAPNGPAGGAIATQSSSSSR-----PPTRARG-----APSP----- 66  
 121 PSVARSNSLRISATVSPVNVNSARHSRPPTLPVSQKGYFPNDPSYAPFLRNQKPPMS 180

67 ---GVLAGPH-----ASEPOLAPACTPAAPAVNGPPGPPSPREP---ORVS 107  
 181 TTVEVEKPHQYQOITITVAPERTTTPOIQPS-----PSTPQAKQKQPKCTEGVS 230

108 HEQFRALQLVNDVDPSPSYLDNFIKIGEGSTGIVCIATVSSGKLVAAVKMDLRKQQR 167  
 231 DEERNALKFFVNDGTDPRSDLTQKQIGEGSTGVVAAYKISTQOIVAVKMDLRKQQR 290

168 ELFNEVVMADYGHENVVEMKNSYLVGDDELAVVMEFLLEGALTDIYTHMEEQIAAV 227  
 291 ELTFNEVSIIRKQYHPNIVRPFSSHLVDDELAVVMEFLLEGSLTDIYVATMTSPQIATI 350

228 CLAVQLALSVLAAGVHRDIKSDSILLTHDGRVYLSDFGCAOVSKVPRKSLVGTPT 287  
 351 SRQVLGALDPLHAKVHRDIKSDSILLKRDGYKLTDFGCGQLSEVPRRSLVGTPT 410

288 WMABELISRLPYGPEVDIWSIGIVNIEMVDSGPYPFNEBPKAMKMITRDNLPRLKNIHK 347  
 411 WTAEBVIAEBPYDTRADISFGIMLIEVGEGRPPFPNDQPFQAKKRIDEHEAKFSRHAK 470

348 VSPBLKGFRLRLVDRPAQRATTAABLKHPLAKAGPPASIVPLMRQ 394  
 471 VSVLSLSLSHCIVDKVWKWPAKDLRHPFAKAQSSSIAPLLQ 517

## RESULT 3

S58682

protein kinase, p21-activated (EC 2.7.1.1) - human

N/Alternate names: protein kinase PAK65; S6/H4 kinase

C/Species: *Homo sapiens* (man)

C/Date: 22-Aug-1996 #sequence\_revision 06-Sep-1996 #text\_change 16-Aug-2004

C/Accession: S58682; S55304; S58690; A57441

R/Seller, M.; Krause, U.J.; Bagrodia, S.; Ambrose, D.; Bokoch, G.M.; Chernoff, J.

Submitted to the EMBL Data Library, April 1995

A/Description: Human p21-activated protein kinases regulate actin organization in mamma

A/Reference number: S58682

A/Accession: S58682

A/Molecule type: DNA

A/Residues: 1525 &lt;SLE&gt;

A/Cross-references: UNIPROT:Q13177; EMBL:U24153; NID:9780807; PIDN:AAA65442.1; PID:97808

R/Martin, G.A.; Bollag, G.; McCormick, F.; Abo, A.

EMBO J. 14, 1970-1978, 1995

A/Title: A novel serine kinase activated by rac1/CDCA42Hs-dependent autophosphorylation  
 A/Reference number: S55258; MUID:95262637; PMID:7744004  
 A/Accession: S55258

A/Status: not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 1525 &lt;MEETQKSNLEL&gt;, 31-89, 'T', 91-149, 'F', 151-224, 'T', 226-328, 'R', 330-338, 340-525

A/Accession: S55304

A/Molecule type: protein

A/Residues: 402-418 &lt;MAW&gt;

R/Martin, G.A.; Bollag, G.; McCormick, F.; Abo, A.

EMBO J. 14, 4385, 1995

A/Reference number: S58690; MUID:96016211; PMID:7556080

A/Contents: erratum

A/Accession: S58690

A/Status: not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 1-30 &lt;MAP&gt;

R/Banner, G.E.; Dennis, P.B.; Maasacchia, R.A.

J. Biol. Chem. 270, 21121-21128, 1995

A/Title: Activation of an S6/H4 kinase (PAK 65) from human placenta by intramolecular an

A/Reference number: A57441; MUID:95403344; PMID:7673144

A/Accession: A57441

A/Molecule type: protein

A/Residues: 197-216/402, 'S', 404-409 &lt;BEN&gt;

A/Experimental source: placenta

C/Superfamily: protein kinase homology

C/Keywords: ATP; autophosphorylation; phosphoprotein; phosphotransferase; serine/threonin

F/247-501/Domain: protein kinase homology &lt;KIN&gt;

F/255-263/Region: protein kinase ATP-binding motif

F/197,402/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 40.2%; Score 833.5; DB 2; Length 525;  
 Best Local Similarity 44.6%; Pred. No. 2e-26;  
 Matches 173; Conservative 69; Mismatches 109; Indels 37; Gaps 6;

39 PSAGGAIPOSSSSSRPPTRARGAPSGVLAHPHA-SEPOLAPACTP----- 85  
 145 PEKQGL-----PSGTPALNAKGTAPAVTTEEDDEETAPVLAIPRDHTKSIYTRS 197

86 AARVPPPPPPR-----SPQREQRVSHQFRALQLVNDPQSPRYLDNFIIGE 136  
 198 VIDPVPAPVDSHVDAKSLDKQKPKMTDEIRMEKRTIVSIGPKKKYTRYETIGQ 257

137 GSTGIVCIATVSSGKLVAAVKMDLRKQKRELLFNEVIMRDYGHENVVEMYSYLVGD 196  
 258 GASGTVTAIDVALGCVALLKQINLQKPKKELLINELIMKELKNINYPFLDSYLVGD 317

197 ELWVMEFLLEGALTDIVTHTR-MNEEQIAVCLAVQLALSVLAAGVHRDIKSDSILL 255  
 318 ELFVMEYLAGSLTDVVTETACWDEAQIAAVCRECQALEFLHANGVHRDIKSDVVL 377

256 THQGRVYLSDFGCAOVSKVPRKSLVGTPTMMAPBLSRLPYGPEVDIWSIGIVTEM 315  
 378 GMEGSVLTDFGFAQITTPQSKSTVGGTPYMMAPVPRKAVGAPVDIWSIGIVTEM 437

316 VDGEPPYFNEBPKAMKMITRDNLPRLKNIHKVSPSKGFLRLVNDPAPQRTAAETLK 375  
 438 VEGEPYLTENPRLALVLIATNGPBLQNPBKLSPIRDLNCLMDVKGKSAKELLQ 497

376 HPFLAKAGPPASIVL-----NRQNR 396  
 498 HPFLAKAPLSLTPLIMAAKAKMSNR 525

## RESULT 4

A57597

beta-p21-activated protein kinase - rat

N/Alternate names: beta-PAK

C/Species: *Rattus norvegicus* (Norway rat)

C/Date: 08-Feb-1996 #sequence\_revision 08-Feb-1996 #text\_change 09-Jul-2004

C/Accession: A57597

R/Manser, B.; Chong, C.; Zhao, Z.S.; Leung, T.; Michael, G.; Hall, C.; Lim, L.

J. Biol. Chem. 270, 25070-25078, 1995

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2005, 09:50:45 ; Search time 94.1678 Seconds

(Without alignments)  
1634.641 Million cell updates/sec

Title: US-10-725-121A-30

Perfect score: 2073

Sequence: 1 ASGAKTLAAGPFTYPRADT.....LAKGPPASIVPLMRQNR 398

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*  
1: geneseqp1980s:\*  
2: geneseqp1980s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2073	100.0	398	2	AAV55941 Human PAK
2	2073	100.0	591	2	AAV55964 Pull leng
3	2073	100.0	591	3	AAV59128 Human ser
4	2073	100.0	591	7	ADG37307 Nuclear f
5	2073	100.0	591	7	ADD89973 Human can
6	2073	100.0	591	7	ADFA5079 Human kin
7	2073	100.0	591	8	ADH23362 Human p21
8	2073	100.0	591	8	ABM82101
9	2073	100.0	620	4	ABG19308
10	2069	99.8	522	8	ABM84642
11	2069	99.8	522	8	ABM84644
12	2069	99.8	522	8	ABM84645
13	2058	99.3	588	8	ADH23359
14	1963	94.7	593	8	ADJ96654
15	1926	92.9	501	7	ADG37309
16	1925	92.9	438	7	ADG37305
17	1917	92.5	438	8	ADG3763
18	1645	79.4	517	8	ABM84648
19	1645	79.4	517	8	ABM84647
20	1645	79.4	517	8	ABM84643
21	1645	79.4	517	8	ABM84646
22	1502	72.5	293	7	ADG15851
23	1351	65.2	719	7	ADJ83007
24	1345	64.9	719	4	AAV5940
25	1340	64.6	719	4	AAV5705

26	1340	64.6	719	7	ADG37451	ADG37451 Nuclear f
27	1340	64.6	719	7	ADP45080	ADP45080 Human kin
28	1340	64.6	719	7	ADJ83008	ADJ83008 Human mam
29	1340	64.6	719	8	ADJ29314	ADJ29314 Human MAR
30	1339	64.6	719	4	AAE02187	AAE02187 Human p21
31	1330	64.2	457	8	ADH42207	ADH42207 Novel hum
32	1330	64.2	457	8	ADH42215	ADH42215 Novel hum
33	1330	64.2	457	8	ADH42211	ADH42211 Novel hum
34	1330	64.2	457	8	ADH42217	ADH42217 Novel hum
35	1330	64.2	457	8	ADH42213	ADH42213 Novel hum
36	1277	61.6	250	3	AAV59129	AAV59129 Human PAK
37	1255.5	60.6	639	8	ADH23364	ADH23364 Fruit fly
38	1252.5	60.4	635	8	ADH23360	ADH23360 Fruit fly
39	1241	59.9	240	4	AAV93297	AAV93297 Human pol
40	1241	59.9	240	7	ADG37303	ADG37303 Nuclear f
41	1241	59.9	240	8	ADL30760	ADL30760 Human pro
42	1205	58.1	338	7	ADL22704	ADL22704 Human dis
43	1205	58.1	547	4	AAV67825	AAV67825 Human p21
44	1197	57.7	681	2	AAV55940	AAV55940 Human PAK
45	1197	57.7	681	3	AAV03967	AAV03967 Signal tr

## ALIGNMENTS

RESULT 1	AAV55941	standard; protein; 398 AA.
ID	AAV55941	
XX	AAV55941:	
AC	18-FEB-2000	(first entry)
XX	Human PAK5 protein.	
XX		
KW	Antithrombotic; antithrombotic; antiinflammatory; antiallergic; osteopathic;	
KW	antiproliferative; antiarteriosclerotic; antisthmatic; immunosuppressive;	
KW	neuroprotective; cardiact; cerebroprotective; cytoskeletal; antidiabetic;	
KW	vulnery; STS20; protein kinase; STK2; STK3; STK4; STK5; STK6; STK7;	
KW	ZC1; ZC2; ZC3; ZC4; KHS2; STU1; STU2; STU3; STU4; STU5; STU6; STU7;	
KW	antibody; gene therapy; rheumatoid arthritis; artherosclerosis; asthma;	
KW	inflammatory bowel disease; Crohn's disease; osteoarthritis; psoriasis;	
KW	rhinitis; autoimmunity; organ transplantation; multiple sclerosis;	
KW	myocardial infarction; cardiovascular disease; stroke; renal failure;	
KW	oxidative stress-related neurodegenerative disorder; Parkinson's disease;	
KW	amyotrophic lateral sclerosis; Leigh syndrome; cancer; cardiomyopathy;	
KW	ischemic disorder; inflammation; diabetes mellitus; fibrosis; mitosis;	
KW	mesangial disorder; growth regulation; wound healing; T cell activation;	
KW	immunosuppressant.	
XX		
OS	Homo sapiens.	
XX		
PN	MO9953036-A2.	
XX		
PD	21-OCT-1999.	
XX		
PF	13-APR-1999; 99MO-US008150.	
XX		
PR	14-APR-1998; 98US-0081784P.	
XX		
PA	(SUGR-) SUGEN INC.	
XX		
PI	Plowman G, Martinez R, Whyte D,	
XX		
DR	WPI, 1999-611301/52.	
XX		
PT	N-PADB; AA240493.	
XX		
PT	Novel kinase-related polypeptides used for the diagnosis and treatment of	
XX	kinase-related diseases and disorders.	
PS	Disclosure; Page 310-312; 387pp; English.	
XX		
CC	This sequence represents a novel STE20-related protein kinase. The	

invention relates to nucleic acid molecule encoding a kinase polypeptide selected from STLK1, STLK3, STLK4, STLK5, STLK6, STLK7, ZC1, ZC2, ZC3, ZC4, KHS2, STUUI, STUUX, GEX2, PKA4 and PAK5. The proteins are used to identify agonists and antagonists, and to raise antibodies. The polynucleotides are useful in gene therapy protocols. The polynucleotides, polypeptides, antibodies, antagonists and agonists may be used to treat diseases such as immune-related disorders and diseases (e.g. rheumatoid arthritis, artherosclerosis, chronic inflammatory bowel disease (e.g. Crohn's disease), asthma, osteoarthritis, psoriasis, atherosclerosis, rhinitis, autoimmunity, and organ transplantation, chronic inflammatory pelvic disease, multiple sclerosis, organ transplantation, myocardial infarction, cardiovascular diseases, stroke, renal failure, oxidative stress-related neurodegenerative disorders (e.g. amyotrophic lateral sclerosis, Parkinson's disease and Leigh syndrome), cancer, cardiomyopathies, ischemic disorders, inflammatory disorders, diabetes mellitus, fibrotic and mesangial disorders. The proteins may also be useful for cell growth regulation (e.g. in wound healing), T cell activation, mitosis control, and as immunosuppressants

Query Match	100.0%	Score 2073	DB 2	Length 398
Best Local Similarity	100.0%	Pred. No. 1.9e-157		
Matches 398	0	Mismatches	0	Gaps 0

Qy	1	USGKTLAAGRPNNTYPRADTDPHSAGAGEPHDAVNPNSAGGLAIPOSSSSSSSRPTRA	60
Db	1	ASGALTAAGRPNNTYPRADTDPHSAGAGEPHDAVNPNSAGGLAIPOSSSSSSSRPTRA	60
Qy	61	RGAPSPGVLGPHASBPQLAPACTPAADAVPGPPGRSPQREBPQVSHQEPRAALQLVVD	120
Db	61	RGAPSPGVLGPHASBPQLAPACTPAADAVPGPPGRSPQREBPQVSHQEPRAALQLVVD	120
Qy	121	PGDPRSILDNFKIGEGSTGIVCIATVASSGGLVAVKKMDLRKQQRRELLFNEVIMEDY	180
Db	121	PGDPRSILDNFKIGEGSTGIVCIATVASSGGLVAVKKMDLRKQQRRELLFNEVIMEDY	180
Qy	181	QHENVNVEYNSTYLVDDELMTVMMEFLREGALDTIVHTHNMNEQLNAVCLATVQALSTVHA	240
Db	181	QHENVNVEYNSTYLVDDELMTVMMEFLREGALDTIVHTHNMNEQLNAVCLATVQALSTVHA	240
Qy	241	QGVIRHDKSDSIILLTHDGRVGLSDPGFCAQVSKVEPPRKSIVGTPTVMAPELISRLPYG	300
Db	241	QGVIRHDKSDSIILLTHDGRVGLSDPGFCAQVSKVEPPRKSIVGTPTVMAPELISRLPYG	300
Qy	301	PEVDIWSLGINVIENWVDGEPYPYFNBPPLKAMKMIRDNDLPPRLKNIHLKVSPSLKGFLDBLL	360
Db	301	PEVDIWSLGINVIENWVDGEPYPYFNBPPLKAMKMIRDNDLPPRLKNIHLKVSPSLKGFLDBLL	360
Qy	361	VRDPAQRATTAABELKHPLAKAGPPASIVPLMRQNRTR	398
Db	361	VRDPAQRATTAABELKHPLAKAGPPASIVPLMRQNRTR	398

RESULT 2  
AA55964  
ID AA55964 standard; protein; 591 AA.

XX		
AC	AAY55964,	
XX		
DT	18-FEB-2000	(first entry)
XX		
DE	Full length human PAK5 protein.	

KM Antihypertensive; antiarrhythmic; antiinflammatory; antiallergic; osteopathic;  
KM antipsychotic; antiarteriosclerotic; antiaschematic; immunosuppressive;  
KM neuroprotective; cardiatic; cerebroprotective; cytostatic; antididiabetic;  
KM vulnery; SREB2; protein kinase; SLK3; SLK3; SLK3; SLK3; SLK3;  
KM ZC1; ZC2; ZC4; KHS3; SOUT1; SOUT3; GSK3; PAK4; PAK5; antagonistic;  
KM antibody; gene therapy; rheumatoid arthritis; atherosclerosis; ashma;  
KM inflammatory bowel disease; Crohn's disease; osteoarthritis; psoriasis;  
KM rthritis; autoimmunity; organ transplant disease; multiple sclerosis;  
KM myocardial infarction; cardiovascular disease; stroke; renal failure;

KM oxidative stress-related neurodegenerative disorder; Parkinson's disease;  
KM amyotrophic lateral sclerosis; Leigh syndrome; cancer; cardiomyopathy;  
KM ischemic disorder; inflammation; diabetes mellitus; fibrosis; mitosis;  
KM megalgal disorder; growth regulation; wound healing; T cell activation;  
KM immunosuppressant.

**Homo sapiens.**

AA PN W09953036-A2.

PD 21-OCT-1999.

PF 13-APR-1999; 99WO-US008150.

PR 14-APR-1998; 98US-0081784P.

PA (SUGE-) SUGEN INC.

PI Plowman G, Martinez R, Whyte D;

WPI; 1999-611301/52.

[illegible]

kinase-related diseases and disorders.

PS Claim 11, page 366-368; 387pp; English.

This sequence represents a novel STE20-related protein kinase. The invention relates to nucleic acid molecule encoding a kinase polypeptide selected from STRK2, STRK3, STRK4, STRK5, STRK6, STRK7, ZC1, ZC2, ZC3, ZC4, KHS2, SUIU1, SUIU3, GEX2, PAK5 and PAK5. The proteins are used to identify agonists and antagonists, and to raise antibodies. The polynucleotides are useful in gene therapy protocols. The polynucleotides, polypeptides, antibodies, antagonists and agonists may be used to treat diseases such as immune-related disorders and diseases (e.g. rheumatoid arthritis, atherosclerosis, chronic inflammatory bowel disease (e.g. Crohn's disease), asthma, osteoarthritis, psoriasis, atherosclerosis, thallitis, autoimmunity, and organ transplantation, chronic inflammatory pelvic disease, multiple sclerosis, organ transplantation, myocardial infarction, cardiovascular disease, stroke, renal failure, oxidative stress-related neurodegenerative disorders (e.g. amyotrophic lateral sclerosis, Parkinson's disease and Leigh syndrome), cancer, cardiomyopathies, ischemic disorders, inflammatory disorders, diabetes mellitus, fibrotic and mesangial disorders. The proteins may also be useful for cell growth regulation (e.g. in wound healing), T cell activation, mitosis control, and as immunosuppressants

AA Sequence 591 AA;  
SQ

Query Match	100.0%	Score 2073;	DB 2;	Length 591;
Best Local Similarity	100.0%	Pred. No. 3.2e-157;		
Matches 398; Conservative	0;	Mismatches	0;	Gaps 0;

Qy	1	ASGATLAAGRPNNTYPRADTQHPBRSAGCEHDVA	PNPSASGGLAIPOSSSSSSPPRA	60
Db	194	ASGATLAAGRPNNTYPRADTQHPBRSAGCEHDVA	PNPSASGGLAIPOSSSSSSPPRA	253
Qy	61	RGAPSPGLGPHASEPOLAPACTPAAAVGPPGPRS	POREPORVSHQEPRAALQLVVD	120
Db	254	RGAPSPGLGPHASEPOLAPACTPAAAVGPPGPRS	POREPORVSHQEPRAALQLVVD	313
Qy	121	PEDPSPSYDNFPIKIGEGSTGIVCIATYASSGKL	YAVKMDLRKQORRELLFNEVYIMDY	180
Db	314	PEDPSPSYDNFPIKIGEGSTGIVCIATYASSGKL	YAVKMDLRKQORRELLFNEVYIMDY	373
Qy	161	QHENVVMYNSYLVDGLVMYMEFLLEGALTDI	VYTHTRMNEEQIQAACVLAIVLQALSYLHA	240
Db	374	QHENVVMYNSYLVDGLVMYMEFLLEGALTDI	VYTHTRMNEEQIQAACVLAIVLQALSYLHA	433
Qy	241	QGVHRIIDIKSDSILLTHDGRVYLSDFGCAOVSK	VEPPRKSIVGTPYMAAPDELLSRPLRG	300
Db	434	QGVHRIIDIKSDSILLTHDGRVYLSDFGCAOVSK	VEPPRKSIVGTPYMAAPDELLSRPLRG	493

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: November 17, 2005, 10:04:33 ; Search time 57.9494 Seconds

(without alignments)  
2873.660 Million cell updates/sec

Title: US-10-725-121a-30

Perfect score: 2073  
Sequence: 1 ASGAKLAAGRPNTYPRADT.....LAKAGPPASIVPLMRQNRTR 398Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1867879 seqs, 418409474 residues

Total number of hits satisfying chosen parameters: 1867879

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:\*

2: /cgn2\_6/ptodata/1/pubppaa/PCT\_NEW\_PUB.pep:\*

3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep:\*

4: /cgn2\_6/ptodata/1/pubppaa/US06\_PUBCOMB.pep:\*

5: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep:\*

6: /cgn2\_6/ptodata/1/pubppaa/PCTUS\_PUBCOMB.pep:\*

7: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep:\*

8: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep:\*

9: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*

10: /cgn2\_6/ptodata/1/pubppaa/US09B\_PUBCOMB.pep:\*

11: /cgn2\_6/ptodata/1/pubppaa/US09C\_PUBCOMB.pep:\*

12: /cgn2\_6/ptodata/1/pubppaa/US09\_NEW\_PUB.pep:\*

13: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*

14: /cgn2\_6/ptodata/1/pubppaa/US10B\_PUBCOMB.pep:\*

15: /cgn2\_6/ptodata/1/pubppaa/US10C\_PUBCOMB.pep:\*

16: /cgn2\_6/ptodata/1/pubppaa/US10D\_PUBCOMB.pep:\*

17: /cgn2\_6/ptodata/1/pubppaa/US10E\_PUBCOMB.pep:\*

18: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.pep:\*

19: /cgn2\_6/ptodata/1/pubppaa/US11\_PUBCOMB.pep:\*

20: /cgn2\_6/ptodata/1/pubppaa/US11\_NEW\_PUB.pep:\*

21: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep:\*

22: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2073	100.0	398	10	US-09-291-417-30
2	2073	100.0	398	16	US-10-725-329-30
3	2073	100.0	398	18	US-10-725-121-30
4	2073	100.0	591	10	US-09-291-417-103
5	2073	100.0	591	14	US-10-134-102-4
6	2073	100.0	591	15	US-10-394-322A-48
7	2073	100.0	591	15	US-10-693-367-2
8	2073	100.0	591	16	US-10-725-329-103
9	2073	100.0	591	17	US-10-893-025-5
10	2073	100.0	591	18	US-10-725-121-103
11	2073	100.0	591	18	US-10-509-599-18

12	2073	100.0	620	18	US-10-450-763-49667	Sequence 49667, A
13	2058	99.3	588	14	US-10-134-102-1	Sequence 1, App11
14	1963	94.7	593	16	US-10-618-941-111	Sequence 111, App
15	1502	72.5	292	15	US-10-406-676-5	Sequence 5, App11
16	1502	72.5	292	15	US-10-406-676-6	Sequence 6, App11
17	1502	72.5	292	15	US-10-406-676-8	Sequence 8, App11
18	1502	72.5	292	15	US-10-406-676-9	Sequence 9, App11
19	1502	72.5	292	15	US-10-406-676-10	Sequence 10, App1
20	1502	72.5	292	15	US-10-406-676-11	Sequence 11, App1
21	1502	72.5	293	15	US-10-406-676-4	Sequence 4, App11
22	1481	71.4	292	15	US-10-406-676-7	Sequence 7, App11
23	1481	71.4	292	15	US-10-406-676-12	Sequence 12, App1
24	1421	68.5	292	15	US-10-406-676-15	Sequence 15, App1
25	1351	65.2	719	14	US-10-331-095-2	Sequence 2, App11
26	1340	64.6	719	14	US-10-331-095-4	Sequence 4, App11
27	1340	64.6	719	15	US-10-394-322A-49	Sequence 49, App1
28	1303	62.9	290	15	US-10-406-676-13	Sequence 13, App1
29	1289	62.2	290	15	US-10-406-676-14	Sequence 14, App1
30	1289	62.2	290	15	US-10-406-676-16	Sequence 16, App1
31	1289	62.2	290	15	US-10-406-676-17	Sequence 17, App1
32	1277	61.6	250	15	US-10-693-367-3	Sequence 3, App11
33	1255.5	60.6	635	14	US-10-134-102-6	Sequence 6, App11
34	1252.5	60.4	635	14	US-10-134-102-2	Sequence 2, App11
35	1213	58.5	290	15	US-10-406-676-18	Sequence 18, App1
36	1209	58.3	290	15	US-10-406-676-19	Sequence 19, App1
37	1209	58.3	290	15	US-10-406-676-20	Sequence 20, App1
38	1197	57.7	681	10	US-09-291-417-29	Sequence 29, App1
39	1197	57.7	681	16	US-10-725-329-29	Sequence 29, App1
40	1197	57.7	681	16	US-10-737-450-108	Sequence 108, App
41	1197	57.7	681	16	US-10-885-921-10	Sequence 10, App1
42	1197	57.7	681	18	US-10-725-121-29	Sequence 29, App1
43	1194	57.6	681	10	US-09-765-815-2	Sequence 2, App11
44	1186.5	57.2	682	18	US-10-840-512-209	Sequence 209, App
45	1147	55.3	292	15	US-10-406-676-21	Sequence 21, App1

## ALIGNMENTS

RESULT 1  
US-09-291-417-30  
; Sequence 30, Application US/09291417A  
; Publication No. US20030050230A1  
; GENERAL INFORMATION:  
; APPLICANT: PLOMMAN, GREGORY  
; APPLICANT: MARTINEZ, RICARDO  
; APPLICANT: MAYHE, DAVID  
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASRS  
; FILE REFERENCE: 240/300  
; CURRENT APPLICATION NUMBER: US/09/291,417A  
; CURRENT FILING DATE: 1999-04-13  
; EARLIER APPLICATION NUMBER: US 60/081,784  
; EARLIER FILING DATE: 1998-04-14  
; NUMBER OF SEQ. ID NOS: 147  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 30  
; LENGTH: 398  
; TYPE: PRT  
; ORGANISM: Mammalian (Human) PAK5  
US-09-291-417-30

Query Match 100.0%; Score 2073; DB 10; Length 398;  
Best Local Similarity 100.0%; Pred. No. 1.4e-128;  
Matches 398; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ASGAKLAAGRPNTYPRADTDPHSGAGSPHYAVPGSPAGGIATPOSSSSSRPPTRA	60
DB	1	ASGAKLAAGRPNTYPRADTDPHSGAGSPHYAVPGSPAGGIATPOSSSSSRPPTRA	60
QY	61	RGAPSPGVLPSPHSEPOLAPACTPAAPAVGPGPSPQREBPQVSHQGFRAALQLVVD	120
DB	61	RGAPSPGVLPSPHSEPOLAPACTPAAPAVGPGPSPQREBPQVSHQGFRAALQLVVD	120

QY 121 PGDPRSYLDNFPIKIGESTGIVCIATVRSRSGKLVAVKMDLRKQORRELLFNEVVIWMDY 180  
DB 121 PGDPRSYLDNFPIKIGESTGIVCIATVRSRSGKLVAVKMDLRKQORRELLFNEVVIWMDY 180  
QY 181 QHENVVMYNSYLVGDELMVMEFLBEGGALTDIVTHTRMNEBOJAAVCLAVLQALSVLHA 240  
DB 181 QHENVVMYNSYLVGDELMVMEFLBEGGALTDIVTHTRMNEBOJAAVCLAVLQALSVLHA 240  
QY 241 QGVTHRDIKSDSILLTHDGRVKLSDFGFCQAVSKEVPRRSKSLVGTPTWMABELISRLLPYG 300  
DB 241 QGVTHRDIKSDSILLTHDGRVKLSDFGFCQAVSKEVPRRSKSLVGTPTWMABELISRLLPYG 300  
QY 301 PEVDIWSLGIMVLEMVNDGEPPEPYFNEPPLKAMKMRDNLPRLLKNLHKVSPSLKGFLLDRLL 360  
DB 301 PEVDIWSLGIMVLEMVNDGEPPEPYFNEPPLKAMKMRDNLPRLLKNLHKVSPSLKGFLLDRLL 360  
QY 361 VRDPAQRATAAEELKHPEFLAKAGPPASIVPLMRQNRTR 398  
DB 361 VRDPAQRATAAEELKHPEFLAKAGPPASIVPLMRQNRTR 398

## RESULT 2

US-10-725-329-30  
Sequence 30, Application US/10725329  
Publication No. US20040224323A1  
GENERAL INFORMATION:  
APPLICANT: PLOMMAN, GREGORY  
APPLICANT: MARTINEZ, RICARDO  
APPLICANT: WHITE, DAVID  
TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES  
FILE REFERENCE: 038602/0328  
CURRENT APPLICATION NUMBER: US/10/725,329  
CURRENT FILING DATE: 2003-12-02  
PRIOR APPLICATION NUMBER: US/09/688,188B  
PRIOR FILING DATE: 2000-10-16  
PRIOR APPLICATION NUMBER: 09/291,417  
PRIOR FILING DATE: 1999-04-14  
PRIOR APPLICATION NUMBER: 60/081,784  
PRIOR FILING DATE: 1998-04-14  
NUMBER OF SEQ ID NOS: 155  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 30  
LENGTH: 398  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-725-329-30

Query Match 100.0%; Score 2073; DB 16; Length 398;  
Best Local Similarity 100.0%; Pred. No. 1,4e-128;

Matches 398; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASGAKLAAGRPFNTYPPRADTDHPSRGAQGBPHDVAFNGPSAGGLAIPQSSSSSRPPTRA 60  
DB 1 ASGAKLAAGRPFNTYPPRADTDHPSRGAQGBPHDVAFNGPSAGGLAIPQSSSSSRPPTRA 60  
QY 61 RGAPSPGVLGPHASEBPOLAPACTPAAPAVGPPGPPSPQREPORVSHBOFRALQLVVD 120  
DB 61 RGAPSPGVLGPHASEBPOLAPACTPAAPAVGPPGPPSPQREPORVSHBOFRALQLVVD 120  
QY 121 PGDPRSYLDNFPIKIGESTGIVCIATVRSRSGKLVAVKMDLRKQORRELLFNEVVIWMDY 180  
DB 121 PGDPRSYLDNFPIKIGESTGIVCIATVRSRSGKLVAVKMDLRKQORRELLFNEVVIWMDY 180  
QY 181 QHENVVMYNSYLVGDELMVMEFLBEGGALTDIVTHTRMNEBOJAAVCLAVLQALSVLHA 240  
DB 181 QHENVVMYNSYLVGDELMVMEFLBEGGALTDIVTHTRMNEBOJAAVCLAVLQALSVLHA 240  
QY 241 QGVTHRDIKSDSILLTHDGRVKLSDFGFCQAVSKEVPRRSKSLVGTPTWMABELISRLLPYG 300  
DB 241 QGVTHRDIKSDSILLTHDGRVKLSDFGFCQAVSKEVPRRSKSLVGTPTWMABELISRLLPYG 300  
QY 301 PEVDIWSLGIMVLEMVNDGEPPEPYFNEPPLKAMKMRDNLPRLLKNLHKVSPSLKGFLLDRLL 360  
DB 301 PEVDIWSLGIMVLEMVNDGEPPEPYFNEPPLKAMKMRDNLPRLLKNLHKVSPSLKGFLLDRLL 360

DB 301 PEVDIWSLGIMVLEMVNDGEPPEPYFNEPPLKAMKMRDNLPRLLKNLHKVSPSLKGFLLDRLL 360  
QY 361 VRDPAQRATAAEELKHPEFLAKAGPPASIVPLMRQNRTR 398  
DB 361 VRDPAQRATAAEELKHPEFLAKAGPPASIVPLMRQNRTR 398

## RESULT 3

US-10-725-121-30  
Sequence 30, Application US/10725121  
Publication No. US20050142625A1  
GENERAL INFORMATION:  
APPLICANT: PLOMMAN, GREGORY  
APPLICANT: MARTINEZ, RICARDO  
APPLICANT: WHITE, DAVID  
TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES  
FILE REFERENCE: 038602/0328  
CURRENT APPLICATION NUMBER: US/10/725,121  
CURRENT FILING DATE: 2003-12-02  
PRIOR APPLICATION NUMBER: US/09/688,188B  
PRIOR FILING DATE: 2000-10-16  
PRIOR APPLICATION NUMBER: 09/291,417  
PRIOR FILING DATE: 1999-04-14  
PRIOR APPLICATION NUMBER: 60/081,784  
PRIOR FILING DATE: 1998-04-14  
NUMBER OF SEQ ID NOS: 155  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 30  
LENGTH: 398  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-725-121-30

Query Match 100.0%; Score 2073; DB 18; Length 398;  
Best Local Similarity 100.0%; Pred. No. 1,4e-128;  
Matches 398; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASGAKLAAGRPFNTYPPRADTDHPSRGAQGBPHDVAFNGPSAGGLAIPQSSSSSRPPTRA 60  
DB 1 ASGAKLAAGRPFNTYPPRADTDHPSRGAQGBPHDVAFNGPSAGGLAIPQSSSSSRPPTRA 60  
QY 61 RGAPSPGVLGPHASEBPOLAPACTPAAPAVGPPGPPSPQREPORVSHBOFRALQLVVD 120  
DB 61 RGAPSPGVLGPHASEBPOLAPACTPAAPAVGPPGPPSPQREPORVSHBOFRALQLVVD 120  
QY 121 PGDPRSYLDNFPIKIGESTGIVCIATVRSRSGKLVAVKMDLRKQORRELLFNEVVIWMDY 180  
DB 121 PGDPRSYLDNFPIKIGESTGIVCIATVRSRSGKLVAVKMDLRKQORRELLFNEVVIWMDY 180  
QY 181 QHENVVMYNSYLVGDELMVMEFLBEGGALTDIVTHTRMNEBOJAAVCLAVLQALSVLHA 240  
DB 181 QHENVVMYNSYLVGDELMVMEFLBEGGALTDIVTHTRMNEBOJAAVCLAVLQALSVLHA 240  
QY 241 QGVTHRDIKSDSILLTHDGRVKLSDFGFCQAVSKEVPRRSKSLVGTPTWMABELISRLLPYG 300  
DB 241 QGVTHRDIKSDSILLTHDGRVKLSDFGFCQAVSKEVPRRSKSLVGTPTWMABELISRLLPYG 300  
QY 301 PEVDIWSLGIMVLEMVNDGEPPEPYFNEPPLKAMKMRDNLPRLLKNLHKVSPSLKGFLLDRLL 360  
DB 301 PEVDIWSLGIMVLEMVNDGEPPEPYFNEPPLKAMKMRDNLPRLLKNLHKVSPSLKGFLLDRLL 360  
QY 361 VRDPAQRATAAEELKHPEFLAKAGPPASIVPLMRQNRTR 398  
DB 361 VRDPAQRATAAEELKHPEFLAKAGPPASIVPLMRQNRTR 398

## RESULT 4

US-09-291-417-103  
Sequence 103, Application US/09291417A  
Publication No. US20030050230A1  
GENERAL INFORMATION:  
APPLICANT: PLOMMAN, GREGORY  
APPLICANT: MARTINEZ, RICARDO

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: November 17, 2005, 09:51:36 ; Search time 26.5602 Seconds  
(without alignments)  
1118.605 Million cell updates/sec

Title: US-10-725-121a-30

Perfect score: 2073  
Sequence: 1 ASGAKLAAGRPNTYPRADT.....LAKAGPASTVPLMRQNR 398

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

## Database :

Issued Patents AA: \*  
1: /cgn2\_6/ptodaca/1/1aa/5A COMB pep: \*  
2: /cgn2\_6/ptodaca/1/1aa/5B COMB pep: \*  
3: /cgn2\_6/ptodaca/1/1aa/6A COMB pep: \*  
4: /cgn2\_6/ptodaca/1/1aa/6B COMB pep: \*  
5: /cgn2\_6/ptodaca/1/1aa/PCUTUS COMB pep: \*  
6: /cgn2\_6/ptodaca/1/1aa/backfile1 pep: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	* Match Length	DB	ID	Description
1	2073	100.0	398	US-09-688-188B-30	Sequence 30, Appl
2	2073	100.0	398	US-09-291-417D-30	Sequence 30, Appl
3	2073	100.0	591	US-09-082-737-2	Sequence 2, Appl
4	2073	100.0	591	US-09-688-188B-103	Sequence 103, App
5	2073	100.0	591	US-09-718-032-2	Sequence 2, Appl
6	2073	100.0	591	US-09-291-417D-103	Sequence 103, App
7	2073	100.0	591	US-09-949-016-6665	Sequence 6665, App
8	2073	100.0	620	US-09-949-016-7206	Sequence 7206, App
9	1277	61.6	250	US-09-082-737-3	Sequence 3, Appl
10	1277	61.6	250	US-09-718-032-3	Sequence 3, Appl
11	1197	57.7	681	US-09-688-188B-29	Sequence 29, Appl
12	1197	57.7	681	US-09-291-417D-29	Sequence 29, Appl
13	1197	57.7	694	US-09-949-016-10445	Sequence 10445, A
14	1194	57.6	681	US-09-765-815-2	Sequence 2, Appl
15	843	40.7	544	US-09-949-016-11562	Sequence 11562, A
16	842	40.6	524	US-09-538-092-1301	Sequence 1301, Ap
17	833.5	40.2	465	US-08-114-555A-2	Sequence 2, Appl
18	833.5	40.2	506	US-08-369-780-2	Sequence 2, Appl
19	833.5	40.2	506	US-08-475-682-2	Sequence 2, Appl
20	833.5	40.2	506	US-08-780-833-2	Sequence 2, Appl
21	833.5	40.2	506	US-08-636-036-2	Sequence 2, Appl
22	833.5	40.2	506	US-08-918-509-2	Sequence 2, Appl
23	833.5	40.2	506	US-09-108-262-2	Sequence 2, Appl
24	833.5	40.2	506	US-09-688-188B-94	Sequence 94, Appl
25	833.5	39.9	544	US-09-291-417D-94	Sequence 94, Appl
26	827.5	39.9	544	US-08-935-760-2	Sequence 2, Appl
27	825	39.8	524	US-08-615-942A-2	Sequence 2, Appl

28	825	39.8	524	US-09-237-325-2	Sequence 2, Appl
29	820.5	39.6	544	US-08-559-397A-19	Sequence 19, Appl
30	818.5	39.5	551	US-09-949-016-10951	Sequence 10951, A
31	817.5	39.4	545	US-08-935-760-4	Sequence 4, Appl
32	817.5	39.4	545	US-09-688-188B-93	Sequence 93, Appl
33	817.5	39.4	545	US-09-291-417D-93	Sequence 93, Appl
34	817.5	39.4	545	US-09-538-092-1297	Sequence 1297, Ap
35	815.5	39.3	544	US-09-688-188B-95	Sequence 95, Appl
36	815.5	39.3	544	US-09-291-417D-95	Sequence 95, Appl
37	814.5	39.3	465	US-08-559-397A-2	Sequence 2, Appl
38	804	38.8	305	US-09-765-815-10	Sequence 10, Appl
39	793.5	38.3	544	US-08-559-397A-29	Sequence 29, Appl
40	785	37.9	694	US-08-559-397A-31	Sequence 31, Appl
41	779	37.6	544	US-08-559-397A-30	Sequence 30, Appl
42	769	37.1	268	US-08-852-743-3	Sequence 3, Appl
43	769	37.1	268	US-09-185-270-3	Sequence 3, Appl
44	755	36.4	410	US-09-248-796A-18461	Sequence 18461, A
45	741	35.7	250	US-09-718-032-4	Sequence 4, Appl

## ALIGNMENTS

RESULT 1					
US-09-688-188B-30					
Sequence 30, Application US/09688188B					
Patent No. 6656716					
GENERAL INFORMATION:					
APPLICANT: PLOMMAN, GREGORY					
APPLICANT: MARTINEZ, RICARDO					
APPLICANT: WHYTE, DAVID					
TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES					
FILE REFERENCE: 038602/0328					
CURRENT APPLICATION NUMBER: US/09/688,188B					
PRIORITY FILING DATE: 2000-10-16					
PRIOR APPLICATION NUMBER: 09/291,417					
PRIOR FILING DATE: 1999-04-14					
PRIOR APPLICATION NUMBER: 60/081,784					
PRIOR FILING DATE: 1998-04-14					
NUMBER OF SEQ ID NOS: 155					
SOFTWARE: PatentIn Ver. 2.1					
SEQ ID NO 30					
LENGTH: 398					
TYPE: PRT					
ORGANISM: Homo sapiens					
US-09-688-188B-30					
Query Match					
Best Local Similarity 100.0%; Score 2073; DB 4; Length 398;					
Pred. No. 1.8e-155;					
Matches 398; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	ASGAKLAAGRPNTYPRADTDRHPSRGAQGEPHDVAHPNGPSAGGLATPQSSSSSRPPTRA	60		
DB	1	ASGAKLAAGRPNTYPRADTDRHPSRGAQGEPHDVAHPNGPSAGGLATPQSSSSSRPPTRA	60		
QY	61	RGAPSPGVGLAPHASEPQLAPACTPAAPVPGPPRSPQREPVSHQFPALQLVVD	120		
DB	61	RGAPSPGVGLAPHASEPQLAPACTPAAPVPGPPRSPQREPVSHQFPALQLVVD	120		
QY	121	PGSPRYLNFITKIGSGTIVCIATVRSRSGKVAVKQMDLRQQRRELLFNVVIMRDY	180		
DB	121	PGSPRYLNFITKIGSGTIVCIATVRSRSGKVAVKQMDLRQQRRELLFNVVIMRDY	180		
QY	181	QHENVVEMTNSYLVGDELWVMEFLGALTDIVTTRMNEBQIAVCLAVLQALSTLHA	240		
DB	181	QHENVVEMTNSYLVGDELWVMEFLGALTDIVTTRMNEBQIAVCLAVLQALSTLHA	240		
QY	241	QGVTHRDIDKSDSLTLTHDGRVKLSDFGCAQVSKVPRRKSIVGTPTWMAPELISRLPYG	300		
DB	241	QGVTHRDIDKSDSLTLTHDGRVKLSDFGCAQVSKVPRRKSIVGTPTWMAPELISRLPYG	300		
QY	301	PEVDISLIGIMVEMWDGPPYFNEBPLKMKMIRDNLPRLKNLKVSPSLKGFDRLL	360		
DB	301	PEVDISLIGIMVEMWDGPPYFNEBPLKMKMIRDNLPRLKNLKVSPSLKGFDRLL	360		

Db 301 PEVDWISLGIMVEMVDGEPYFNEPPLKAMKMIKMDLPPRLKXNLHKVSPSLKGFIDRL 360  
Qy 361 VRDPAQRATAAEILKHPFLAKAGPPASIVPLMRQNRTR 398  
Db 361 VRDPAQRATAAEILKHPFLAKAGPPASIVPLMRQNRTR 398

RESULT 2

US-09-291-417D-30  
; Sequence 30, Application US/09291417D  
; Patent No. 6680170  
; GENERAL INFORMATION:  
; APPLICANT: PLOMMAN, GREGORY  
; APPLICANT: MARTINEZ, RICARDO  
; APPLICANT: WHITE, DAVID  
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES  
; FILE REFERENCE: 038602/0329  
; CURRENT APPLICATION NUMBER: US/09/291,417D  
; CURRENT FILING DATE: 1999-04-13  
; PRIOR APPLICATION NUMBER: 60/081,784  
; PRIOR FILING DATE: 1998-04-14  
; NUMBER OF SEQ ID NOS: 155  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 30  
; LENGTH: 398  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-291-417D-30

Query Match 100.0%; Score 2073; DB 4; Length 398;  
Best Local Similarity 100.0%; Pred. No. 1.8e-155;  
Matches 398; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASGAKLAAGRPNTYPPADTDHPSRGAQGEPRHDVAENGPSAGGLATPQSSSSSRPPTRA 60  
Db 1 ASGAKLAAGRPNTYPPADTDHPSRGAQGEPRHDVAENGPSAGGLATPQSSSSSRPPTRA 60  
Qy 61 RGAPSPGVLPNHAASEPOLAPACTPPAAYVGPGRPSPOREPORVHEQFRALQLVND 120  
Db 61 RGAPSPGVLPNHAASEPOLAPACTPPAAYVGPGRPSPOREPORVHEQFRALQLVND 120  
Qy 121 PGDPRSYLDNFIKIGESTGIVCIATVRSSGKLVAVKMDLRKQORRELLFNEVIMRDY 180  
Db 121 PGDPRSYLDNFIKIGESTGIVCIATVRSSGKLVAVKMDLRKQORRELLFNEVIMRDY 180  
Qy 181 QHENVEMTNSYLVGDELMVWMEFLBEGALTDIVTHTRMNEBOIAVCLAVLQALSTVLA 240  
Db 181 QHENVEMTNSYLVGDELMVWMEFLBEGALTDIVTHTRMNEBOIAVCLAVLQALSTVLA 240  
Qy 241 QGVTHRIKSDSILLTHDGRVKLSDFGCAQVSKEVPRKSLVGTPTWMAPELISRLPYG 300  
Db 241 QGVTHRIKSDSILLTHDGRVKLSDFGCAQVSKEVPRKSLVGTPTWMAPELISRLPYG 300  
Qy 301 PEVDWISLGIMVEMVDGEPYFNEPPLKAMKMIKMDLPPRLKXNLHKVSPSLKGFIDRL 360  
Db 301 PEVDWISLGIMVEMVDGEPYFNEPPLKAMKMIKMDLPPRLKXNLHKVSPSLKGFIDRL 360  
Qy 361 VRDPAQRATAAEILKHPFLAKAGPPASIVPLMRQNRTR 398  
Db 361 VRDPAQRATAAEILKHPFLAKAGPPASIVPLMRQNRTR 398

RESULT 3

US-09-082-737-2  
; Sequence 2, Application US/09082737  
; Patent No. 6013500  
; GENERAL INFORMATION:  
; APPLICANT: Minden, Audrey  
; APPLICANT: PAK4, A No. 6013500e1 Gene Encoding A Serine/  
; TITLE OF INVENTION: Threonine Kinase  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSES: Cooper & Dunham LLP

STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 11230

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/082,737  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 0575/55311  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 591 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-082-737-2

Query Match 100.0%; Score 2073; DB 3; Length 591;  
Best Local Similarity 100.0%; Pred. No. 3e-155;  
Matches 398; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASGAKLAAGRPNTYPPADTDHPSRGAQGEPRHDVAENGPSAGGLATPQSSSSSRPPTRA 60  
Db 194 ASGAKLAAGRPNTYPPADTDHPSRGAQGEPRHDVAENGPSAGGLATPQSSSSSRPPTRA 253  
Qy 61 RGAPSPGVLPNHAASEPOLAPACTPPAAYVGPGRPSPOREPORVHEQFRALQLVND 120  
Db 254 RGAPSPGVLPNHAASEPOLAPACTPPAAYVGPGRPSPOREPORVHEQFRALQLVND 313  
Qy 121 PGDPRSYLDNFIKIGESTGIVCIATVRSSGKLVAVKMDLRKQORRELLFNEVIMRDY 180  
Db 314 PGDPRSYLDNFIKIGESTGIVCIATVRSSGKLVAVKMDLRKQORRELLFNEVIMRDY 373  
Qy 181 QHENVEMTNSYLVGDELMVWMEFLBEGALTDIVTHTRMNEBOIAVCLAVLQALSTVLA 240  
Db 374 QHENVEMTNSYLVGDELMVWMEFLBEGALTDIVTHTRMNEBOIAVCLAVLQALSTVLA 433  
Qy 241 QGVTHRIKSDSILLTHDGRVKLSDFGCAQVSKEVPRKSLVGTPTWMAPELISRLPYG 300  
Db 434 QGVTHRIKSDSILLTHDGRVKLSDFGCAQVSKEVPRKSLVGTPTWMAPELISRLPYG 493  
Qy 301 PEVDWISLGIMVEMVDGEPYFNEPPLKAMKMIKMDLPPRLKXNLHKVSPSLKGFIDRL 360  
Db 494 PEVDWISLGIMVEMVDGEPYFNEPPLKAMKMIKMDLPPRLKXNLHKVSPSLKGFIDRL 553  
Qy 361 VRDPAQRATAAEILKHPFLAKAGPPASIVPLMRQNRTR 398  
Db 554 VRDPAQRATAAEILKHPFLAKAGPPASIVPLMRQNRTR 591

RESULT 4

US-09-688-188B-103  
; Sequence 103, Application US/09688188B  
; Patent No. 6656716  
; GENERAL INFORMATION:  
; APPLICANT: PLOMMAN, GREGORY  
; APPLICANT: MARTINEZ, RICARDO  
; APPLICANT: WHITE, DAVID  
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES  
; FILE REFERENCE: 038602/0328  
; CURRENT APPLICATION NUMBER: US/09/688,188B



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2005, 10:04:33 ; Search time 86.0506 Seconds  
(without alignments)  
2873.660 Million cell updates/sec

Title: US-10-725-121A-103  
Perfect score: 3090  
Sequence: 1 MFGRRKRRVISAAPSNEHR.....LAKGPPASTVPLMRQNR 591

Scoring table: BIOSOM62  
Gapop 10.0 , Gapext 0.5

Searched: 1867879 seqs, 418409474 residues  
Total number of hits satisfying chosen parameters: 1867879

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: Published Applications AA: \*  
2: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09C\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10F\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US10G\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3090	100.0	591	10	US-09-291-417-103
2	3090	100.0	591	14	US-10-134-102-4
3	3090	100.0	591	15	US-10-394-322A-48
4	3090	100.0	591	15	US-10-693-367-2
5	3090	100.0	591	16	US-10-725-329-103
6	3090	100.0	591	17	US-10-893-025-5
7	3090	100.0	591	18	US-10-725-121-103
8	3090	100.0	591	18	US-10-509-599-18
9	3090	100.0	620	18	US-10-450-763-49667
10	3075	99.5	588	14	US-10-134-102-1
11	2860	92.6	593	16	US-10-618-941-111

12	2073	67.1	398	10	US-09-291-417-30	Sequence 30, App1
13	2073	67.1	398	16	US-10-725-329-30	Sequence 30, App1
14	2073	67.1	398	18	US-10-725-121-30	Sequence 20, App1
15	1748	56.6	719	14	US-10-331-095-2	Sequence 4, App1
16	1731	56.0	719	15	US-10-331-095-4	Sequence 4, App1
17	1731	56.0	719	15	US-10-394-322A-49	Sequence 49, App1
18	1502	48.6	292	15	US-10-406-676-5	Sequence 5, App1
19	1502	48.6	292	15	US-10-406-676-6	Sequence 6, App1
20	1502	48.6	292	15	US-10-406-676-8	Sequence 8, App1
21	1502	48.6	292	15	US-10-406-676-9	Sequence 9, App1
22	1502	48.6	292	15	US-10-406-676-10	Sequence 10, App1
23	1502	48.6	292	15	US-10-406-676-11	Sequence 11, App1
24	1502	48.6	292	15	US-10-406-676-12	Sequence 12, App1
25	1497.5	48.5	639	14	US-10-134-102-6	Sequence 6, App1
26	1494.5	48.4	635	14	US-10-134-102-2	Sequence 2, App1
27	1481	47.9	292	15	US-10-406-676-7	Sequence 7, App1
28	1481	47.9	292	15	US-10-406-676-12	Sequence 12, App1
29	1473.5	47.7	681	10	US-09-291-417-29	Sequence 29, App1
30	1473.5	47.7	681	16	US-10-725-329-29	Sequence 29, App1
31	1473.5	47.7	681	16	US-10-737-450-108	Sequence 108, App
32	1473.5	47.7	681	16	US-10-885-921-10	Sequence 10, App1
33	1473.5	47.7	681	18	US-10-725-121-29	Sequence 29, App1
34	1470.5	47.6	681	10	US-09-765-815-2	Sequence 2, App1
35	1467	47.5	682	18	US-10-840-512-209	Sequence 209, App
36	1421	46.0	292	15	US-10-406-676-15	Sequence 15, App1
37	1335.5	43.2	641	15	US-10-311-034-15	Sequence 15, App1
38	1303	42.2	290	15	US-10-406-676-13	Sequence 13, App1
39	1289	41.7	290	15	US-10-406-676-14	Sequence 14, App1
40	1289	41.7	290	15	US-10-406-676-16	Sequence 16, App1
41	1289	41.7	290	15	US-10-406-676-17	Sequence 17, App1
42	1277	41.3	250	15	US-10-693-367-3	Sequence 3, App1
43	1213	39.3	290	15	US-10-406-676-18	Sequence 18, App1
44	1209	39.1	290	15	US-10-406-676-19	Sequence 19, App1
45	1209	39.1	290	15	US-10-406-676-20	Sequence 20, App1

## ALIGNMENTS

RESULT 1  
US-09-291-417-103  
Sequence 103, Application US/09291417A  
Publication No. US20030050230A1  
GENERAL INFORMATION:  
APPLICANT: PLOOMAN, GREGORY  
APPLICANT: MARTINEZ, RICARDO  
TITLE OR INVENTION: STE20-RELATED PROTEIN KINASES  
FILE REFERENCE: 240/300  
CURRENT APPLICATION NUMBER: US/09/291,417A  
CURRENT FILING DATE: 1999-04-13  
EARLIER APPLICATION NUMBER: 60/081,784  
EARLIER FILING DATE: 1998-04-14  
NUMBER OF SEQ ID NOS: 147  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO 103  
LENGTH: 591  
TYPE: PRT  
ORGANISM: Full Length Mammalian (Human) PAKShu  
US-09-291-417-103

Query Match	100.0%	Score 3090;	DB 10;	Length 591;
Best Local Similarity	100.0%;	Pred. No. 7.5e-161;		
Matches 591;	Conservative 0;	Mismatches 0;	Gaps 0;	
QY	1	MFGRRKRRVISAAPSNEHRVHTGPDHOKFTGLPRQWOSTLIESARRKPLVDPACIT	60	
DB	1	MFGRRKRRVISAAPSNEHRVHTGPDHOKFTGLPRQWOSTLIESARRKPLVDPACIT	60	
QY	61	SIOPGAPKTIIVRSKSGAKOGALTLILDEFENMSVTSNSLRSDSPPPARARQENGWPEE	120	
DB	61	SIOPGAPKTIIVRSKSGAKOGALTLILDEFENMSVTSNSLRSDSPPPARARQENGWPEE	120	

QY 121 PATTARGGPGKAGSRGRFAGHSEAGSGGDRRRRAGPEKRPKSSREGSGGPQESSRDKRPL 180  
DB 121 PATTARGGPGKAGSRGRFAGHSEAGSGGDRRRRAGPEKRPKSSREGSGGPQESSRDKRPL 180  
QY 181 SGPVGTGTPGAGLASGAKLAAGRPFTNTYPRADTDHPSRGAQGBPHDVA PNGPSAGGLAIP 240  
DB 181 SGPVGTGTPGAGLASGAKLAAGRPFTNTYPRADTDHPSRGAQGBPHDVA PNGPSAGGLAIP 240  
QY 241 QSSSSSSRPPTRRAGAPSPGVLAGPHASEPOLAPACTPAAPVAVGPPGPSPOREPQVRS 300  
DB 241 QSSSSSSRPPTRRAGAPSPGVLAGPHASEPOLAPACTPAAPVAVGPPGPSPOREPQVRS 300  
QY 301 HEQFRALQLVDPGDRSYLDNFITKIGBSTGIVCIATVRSSGKLVAVKMDLRKQORR 360  
DB 301 HEQFRALQLVDPGDRSYLDNFITKIGBSTGIVCIATVRSSGKLVAVKMDLRKQORR 360  
QY 361 ELTFNEVIMRDYQHEHNVEMTNSYLVDGLMVMMELEGALTDIYTHTRMNEEQIAAV 420  
DB 361 ELTFNEVIMRDYQHEHNVEMTNSYLVDGLMVMMELEGALTDIYTHTRMNEEQIAAV 420  
QY 421 CLAVLQALSVLHAQGVTHRDIKSDSILLTHDGRVKLSDFGCAQVSKVPRRSKLVGTPT 480  
DB 421 CLAVLQALSVLHAQGVTHRDIKSDSILLTHDGRVKLSDFGCAQVSKVPRRSKLVGTPT 480  
QY 481 WMAPELISRLPYGPEVDIWSLGIWIEVNDGEPYFNEPPLKAMKMTLDNLPPRLKNLHK 540  
DB 481 WMAPELISRLPYGPEVDIWSLGIWIEVNDGEPYFNEPPLKAMKMTLDNLPPRLKNLHK 540  
QY 541 VSPBLKGFPLDRLVDRDPAQRATAELKHPFLAKAGPASIIVPLMRQNRTR 591  
DB 541 VSPBLKGFPLDRLVDRDPAQRATAELKHPFLAKAGPASIIVPLMRQNRTR 591

## RESULT 2

US-10-134-102-4  
; Sequence 4, Application US/10134102  
; Publication No. US20030186254A1  
; GENERAL INFORMATION:  
; APPLICANT: Melnick, Michael B.  
; APPLICANT: Moritz, Albrecht  
; APPLICANT: Comb, Michael J.  
; TITLE OF INVENTION: Regulation of HIV-Tat and Nef by the Pak4 kinase and its  
; FILE REFERENCE: CST-176 CIP  
; CURRENT APPLICATION NUMBER: US/10/134,102  
; PRIOR FILING DATE: 2002-04-29  
; PRIOR APPLICATION NUMBER: 09/750,457  
; PRIOR FILING DATE: 2000-12-28  
; PRIOR APPLICATION NUMBER: 60/173,939  
; PRIOR FILING DATE: 1999-12-30  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 591  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-134-102-4

Query Match 100.0%; Score 3090; DB 14; Length 591;  
Best Local Similarity 100.0%; Pred. No. 7,5e-161;  
Matches 591; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MFGRRKKRVEISAPSNFEHRVHTGPDQHEQKFTGLPRQMSLIBSARRPKELVDPACT 60  
DB 1 MFGRRKKRVEISAPSNFEHRVHTGPDQHEQKFTGLPRQMSLIBSARRPKELVDPACT 60  
QY 61 SIQGPARTIVRSGKAGDGLTLLDDEFENSVTRNSLRDSDPPPARARQENGME 120  
DB 61 SIQGPARTIVRSGKAGDGLTLLDDEFENSVTRNSLRDSDPPPARARQENGME 120  
QY 121 PATTARGGPGKAGSRGRFAGHSEAGSGGDRRRRAGPEKRPKSSREGSGGPQESSRDKRPL 180  
DB 121 PATTARGGPGKAGSRGRFAGHSEAGSGGDRRRRAGPEKRPKSSREGSGGPQESSRDKRPL 180  
QY 241 QSSSSSSRPPTRRAGAPSPGVLAGPHASEPOLAPACTPAAPVAVGPPGPSPOREPQVRS 300  
DB 241 QSSSSSSRPPTRRAGAPSPGVLAGPHASEPOLAPACTPAAPVAVGPPGPSPOREPQVRS 300

QY 181 SGPVGTGTPGAGLASGAKLAAGRPFTNTYPRADTDHPSRGAQGBPHDVA PNGPSAGGLAIP 240  
DB 181 SGPVGTGTPGAGLASGAKLAAGRPFTNTYPRADTDHPSRGAQGBPHDVA PNGPSAGGLAIP 240  
QY 241 QSSSSSSRPPTRRAGAPSPGVLAGPHASEPOLAPACTPAAPVAVGPPGPSPOREPQVRS 300  
DB 241 QSSSSSSRPPTRRAGAPSPGVLAGPHASEPOLAPACTPAAPVAVGPPGPSPOREPQVRS 300  
QY 301 HEQFRALQLVDPGDRSYLDNFITKIGBSTGIVCIATVRSSGKLVAVKMDLRKQORR 360  
DB 301 HEQFRALQLVDPGDRSYLDNFITKIGBSTGIVCIATVRSSGKLVAVKMDLRKQORR 360  
QY 361 ELTFNEVIMRDYQHEHNVEMTNSYLVDGLMVMMELEGALTDIYTHTRMNEEQIAAV 420  
DB 361 ELTFNEVIMRDYQHEHNVEMTNSYLVDGLMVMMELEGALTDIYTHTRMNEEQIAAV 420  
QY 421 CLAVLQALSVLHAQGVTHRDIKSDSILLTHDGRVKLSDFGCAQVSKVPRRSKLVGTPT 480  
DB 421 CLAVLQALSVLHAQGVTHRDIKSDSILLTHDGRVKLSDFGCAQVSKVPRRSKLVGTPT 480  
QY 481 WMAPELISRLPYGPEVDIWSLGIWIEVNDGEPYFNEPPLKAMKMTLDNLPPRLKNLHK 540  
DB 481 WMAPELISRLPYGPEVDIWSLGIWIEVNDGEPYFNEPPLKAMKMTLDNLPPRLKNLHK 540  
QY 541 VSPBLKGFPLDRLVDRDPAQRATAELKHPFLAKAGPASIIVPLMRQNRTR 591  
DB 541 VSPBLKGFPLDRLVDRDPAQRATAELKHPFLAKAGPASIIVPLMRQNRTR 591

## RESULT 3

US-10-394-322A-48  
; Sequence 48, Application US/10394322A  
; Publication No. US2003023391A1  
; GENERAL INFORMATION:  
; APPLICANT: SUNESIS PHARMACEUTICALS, INC.  
; APPLICANT: Prescott, John C.  
; TITLE OF INVENTION: IDENTIFICATION OF KINASE INHIBITORS  
; FILE REFERENCE: 39750-0006 US  
; CURRENT APPLICATION NUMBER: US/10/394,322A  
; PRIOR FILING DATE: 2003-03-20  
; PRIOR APPLICATION NUMBER: US 60/366,892  
; PRIOR FILING DATE: 2002-03-21  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: PasteSeq for Windows Version 4.0  
; SEQ ID NO 48  
; LENGTH: 591  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-394-322A-48

Query Match 100.0%; Score 3090; DB 15; Length 591;  
Best Local Similarity 100.0%; Pred. No. 7,5e-161;  
Matches 591; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MFGRRKKRVEISAPSNFEHRVHTGPDQHEQKFTGLPRQMSLIBSARRPKELVDPACT 60  
DB 1 MFGRRKKRVEISAPSNFEHRVHTGPDQHEQKFTGLPRQMSLIBSARRPKELVDPACT 60  
QY 61 SIQGPARTIVRSGKAGDGLTLLDDEFENSVTRNSLRDSDPPPARARQENGME 120  
DB 61 SIQGPARTIVRSGKAGDGLTLLDDEFENSVTRNSLRDSDPPPARARQENGME 120  
QY 121 PATTARGGPGKAGSRGRFAGHSEAGSGGDRRRRAGPEKRPKSSREGSGGPQESSRDKRPL 180  
DB 121 PATTARGGPGKAGSRGRFAGHSEAGSGGDRRRRAGPEKRPKSSREGSGGPQESSRDKRPL 180  
QY 181 SGPVGTGTPGAGLASGAKLAAGRPFTNTYPRADTDHPSRGAQGBPHDVA PNGPSAGGLAIP 240  
DB 181 SGPVGTGTPGAGLASGAKLAAGRPFTNTYPRADTDHPSRGAQGBPHDVA PNGPSAGGLAIP 240  
QY 241 QSSSSSSRPPTRRAGAPSPGVLAGPHASEPOLAPACTPAAPVAVGPPGPSPOREPQVRS 300  
DB 241 QSSSSSSRPPTRRAGAPSPGVLAGPHASEPOLAPACTPAAPVAVGPPGPSPOREPQVRS 300

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2005, 10:09:48 ; Search time 136.844 Seconds  
(without alignments)  
2211.556 Million cell updates/sec

Title: US-10-725-121A-103

Perfect score: 3090  
Sequence: 1 MFGKRRKRRVISAAPSNEPR.....LAKAGPPASIVPLMRQNRTR 591

Scoring table: BLOSUM62  
Gapop 10.0 , Gapept 0.5

Searched: 1612378 seqs, 512079187 residues  
Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: UniProt 03: \*  
1: uniprot\_sprot: \*  
2: uniprot\_trembl: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3090	100.0	591	1	PAK4 HUMAN
2	2860	92.6	593	1	PAK4 MOUSE
3	2856	92.4	593	1	Q6ZPX0
4	2853	92.3	593	2	Q80Z97
5	2558	82.8	501	2	Q8N481
6	2202.5	71.3	438	2	Q8NDE3
7	2202.5	71.3	467	2	Q9UIS8
8	2124.5	68.8	650	2	Q90M62
9	2031	65.7	663	2	Q6GM11
10	1967	63.7	407	2	Q8K0U2
11	1748	56.6	719	2	Q8C0U5
12	1745	56.5	711	2	Q6NV44
13	1740	56.3	719	2	Q8BVD0
14	1731	56.0	719	1	PAK7 HUMAN
15	1731	56.0	719	2	Q8TBS3
16	1730	56.0	719	2	Q8TBS3
17	1501.5	48.6	639	1	PAKM DROME
18	1473.5	47.7	681	1	PAK6 HUMAN
19	1464.5	47.4	558	2	Q70G55
20	1241	40.2	240	2	Q8NCH5
21	1160	37.5	229	2	Q9CST1
22	1096.5	35.5	540	2	Q9U3M1
23	1090.5	35.3	542	2	Q18637
24	948.5	30.7	704	2	Q70D76
25	937.5	30.3	704	2	Q24190
26	937.5	30.3	704	2	Q9V113
27	936	30.3	827	2	Q6IWM3
28	932.5	30.2	704	2	Q24213
29	924	29.9	544	2	Q8K1R6
30	917	29.7	544	1	PAK3 RAT
31	907	29.4	577	2	Q803Z0

32	906.5	29.3	559	1	PAK3 MOUSE	Q61036	mus musculus
33	905	29.3	577	2	Q6P0I7	Q6P0I7	brachydanio
34	904	29.3	524	1	PAK2 HUMAN	Q13177	homo sapien
35	903	29.2	564	2	Q8AXB4	Q8AXB4	xenopus lae
36	900	29.1	524	1	PAK2 RABIT	Q29502	oryzolaquus
37	895	29.0	524	1	PAK2 MOUSE	Q8C144	mus musculus
38	895	29.0	524	1	PAK2 RAT	Q64303	rattus norv
39	894	28.9	524	2	Q6ISC3	Q6ISC3	homo sapien
40	894	28.9	544	1	PAK1 RAT	P35465	rattus norv
41	892.5	28.9	545	1	PAK1 HUMAN	Q13153	homo sapien
42	891.5	28.9	559	1	PAK3 HUMAN	Q75914	homo sapien
43	891.5	28.9	559	1	PAK3 PANTH	Q7YQ14	pan troglod
44	891.5	28.9	559	1	PAK3 PONY	Q7YQ13	pango pygma
45	890	28.8	517	2	Q8AW67	Q8AW67	brachydanio

## ALIGNMENTS

RESULT 1  
ID PAK4 HUMAN STANDARD; PRT; 591 AA.  
AC Q96013; Q9BU33;  
DT 16-OCT-2001 (Ref. 40, Created)  
DT 16-OCT-2001 (Ref. 40, Last sequence update)  
DT 05-JUL-2004 (Ref. 44, Last annotation update)  
DE Serine/threonine-protein kinase PAK 4 (EC 2.7.1.37) (p21-activated kinase 4) (PAK-4).  
GN Name=PAK4;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RX MEDLINE=99043860; PubMed=982598; DOI=10.1093/emboj/17.22.6527;  
RA Abo A., Qu J., Cammarano M.S., Dan C., Fritsch A., Baud V.,  
RA Belisle B., Minden A.,  
RT "PAK4, a novel effector for Cdc42Hs, is implicated in the reorganization of the actin cytoskeleton and in the formation of filopodia."  
RT EMBO J. 17:6527-6540(1998).  
RL [2]  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Klausner R.D., Collins F.S., Wagner L., Shennan R.D., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marisina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Brownstein M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udell T.B., Tohyuki S., Cavinini P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abrahams R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulys S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green B.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Small D.B.,  
RA Schreuch A., Schein J.E., Jones S.J.M., Marra M.A.,  
RT "generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."  
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -!- FUNCTION: Activates the JNK pathway. Implicated in the reorganization of the actin cytoskeleton and in the formation of filopodia.

```

CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- SUBUNIT: Interacts tightly with GTP-bound but not GPP-bound
CC CDC42/p21 and weakly with RAC1.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named Isoforms=2;
CC Name=1;
CC IsoId=096013-1; Sequence=Displayed;
CC Name=2;
CC IsoId=096013-2; Sequence=VSP_004892, VSP_004893;
CC Note=No experimental confirmation available;
CC TISSUE SPECIFICITY: Highest expression in prostate, testis and
CC colon.
CC -1- PTM: Autophosphorylated when activated by CDC42/P21.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family. STE20
CC subfamily.
CC -1- SIMILARITY: Contains 1 CRIB domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ011855; CA09820.1; -
CC EMBL; AF005046; AAD01210.1; -
CC EMBL; BC002921; AAH02921.1; -
CC EMBL; BC011368; AAH11368.1; -
CC EMBL; BC025282; AAH25282.1; -
CC M5SP; Q13153; 1F3M.
CC Gene; HGNC:16059; PAK4.
CC H-InvDB: HIK0015110; -
CC MIM; 605451; -
CC GO; GO:0005794; C:Golgi apparatus; TAS.
CC GO; GO:0006928; P:cell motility; TAS.
CC GO; GO:0007165; P:signal transduction; TAS.
CC InterPro: IPR011009; Kinase like.
CC InterPro: IPR000095; PAKbox/rho-binding.
CC InterPro: IPR000719; Prot_Kinase.
CC InterPro: IPR008271; Ser_Thr_Pkin_AS.
CC InterPro: IPR002290; Ser_Thr_Pkinase.
CC InterPro: IPR011026; WASP_C.
CC Pfam; PF00706; PBD; 1.
CC Pfam; PF00069; Pkinase; 1.
CC ProDom; PD000001; Prot_Kinase; 1.
CC SMART; SM00285; PBD; 1.
CC SMART; SM00220; S_TKc; 1.
CC PROSITE; PS50108; CRIB; 1.
CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
CC PROSITE; PS50011; PROTEIN KINASE DOM; 1.
CC PROSITE; PS00108; PROTEIN KINASE_ST; FALSE_NEG.
CC Alternative splicing; ATP-binding; Phosphorylation;
CC Serine/threonine-protein kinase; Transferase.
CC FT DOMAIN 11 24
CC FT DOMAIN 25 320
CC FT NP_BIND 321 572
CC FT BINDING 327 335
CC FT ACT_SITE 350 350
CC FT VARSPIC 440 440
CC FT VARSPIC 120 120
CC FT VARSPIC 121 285
CC FT SEQUENCE 591 AA; 64071 MW; 04C2A5C0B06427D5 CRC64;
CC
CC Query Match 100.0%; Score 3090; DB 1; Length 591;
CC Best Local Similarity 100.0%; Pred. No. 3,7e-106;
CC Matches 591; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC 1. MFGRKRRVEISADPSFHRVHTGPDQHKFTGLPRWQSLIESARRPKPLVDPACT 60
CC 1 MFGRKRRVEISADPSFHRVHTGPDQHKFTGLPRWQSLIESARRPKPLVDPACT 60

```

```

QY 61 SIQGAFTIVRSKGAQGLTLTLDPEFNMSTVNSLSRRDPPPARAENGMPBE 120
DB 61 SIQGAFTIVRSKGAQGLTLTLDPEFNMSTVNSLSRRDPPPARAENGMPBE 120
QY 121 PATTARPGPKAGSRGRFAGHSENGSGGSDRRAPGPKRPSRSGSGPQESRDRL 180
DB 121 PATTARPGPKAGSRGRFAGHSENGSGGSDRRAPGPKRPSRSGSGPQESRDRL 180
QY 181 SGPDVGTTPQAGLASGKLAAGRPNTYPRADTDHPSRGAQGEHVDVAPGPSAGLAIP 240
DB 181 SGPDVGTTPQAGLASGKLAAGRPNTYPRADTDHPSRGAQGEHVDVAPGPSAGLAIP 240
QY 241 QSSSSSRPPTRRAGAPSPVGLGPHASEPOLAPACTPAAPVAPGPGSPQSPQSPQ 300
DB 241 QSSSSSRPPTRRAGAPSPVGLGPHASEPOLAPACTPAAPVAPGPGSPQSPQSPQ 300
QY 301 HEQFRALQVDPGDRSYLDNFIKIGSGSTGICVATVRSQGLVAVKMDLRKQQR 360
DB 301 HEQFRALQVDPGDRSYLDNFIKIGSGSTGICVATVRSQGLVAVKMDLRKQQR 360
QY 361 ELFPNRYVTRDQHEVNVEMNNSYVGDLMVYMEELREGALTDYTHTRNMEQIAAV 420
DB 361 ELFPNRYVTRDQHEVNVEMNNSYVGDLMVYMEELREGALTDYTHTRNMEQIAAV 420
QY 421 CLAVIALSVLHAGVTHRDIKSDSILVTHDGRVYKSDPFCQAVSKVPRRKSIVGTPY 480
DB 421 CLAVIALSVLHAGVTHRDIKSDSILVTHDGRVYKSDPFCQAVSKVPRRKSIVGTPY 480
QY 481 WMAPELLSRPLPYGEVDIMSLGIMVIMVDGEPPEYFNEPPLKMKIMRDLPLRLKHLK 540
DB 481 WMAPELLSRPLPYGEVDIMSLGIMVIMVDGEPPEYFNEPPLKMKIMRDLPLRLKHLK 540
QY 541 VSPBLKGFRLRLVDRDPAQRATAELKHPFLAKAGPPASIVPLMRNRR 591
DB 541 VSPBLKGFRLRLVDRDPAQRATAELKHPFLAKAGPPASIVPLMRNRR 591
DB 541 VSPBLKGFRLRLVDRDPAQRATAELKHPFLAKAGPPASIVPLMRNRR 591

RESULT 2
PAK4 MOUSE STANDARD; PRT; 593 AA.
ID PAK4_MOUSE
AC O8BTW9;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Serine/threonine-protein kinase PAK 4 (EC 2.7.1.37) (p21-activated
DE kinase 4) (PAK-4).
GN Name=PAK4;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ND; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schombach C., Gojouri T.,
RA Baldarelli R., Hill D.P., Butt C., Hume D.A., Quackebush J.,
RA Schirml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Choehia C., Corbani L.E., Cousins S.,
RA Dalla B., Dragani T.A., Fletcher C.F., Forrest A., Frazier K.S.,
RA Gaasterland T., Gariboldi M., Glaszi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.U., Jarvis E.D.,
RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA MacIock D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagaishima T., Numata K., Okido T., Pavan W.J., Petras G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: November 17, 2005, 10:00:02 ; Search time 31.6714 Seconds

(Without alignments)  
1795.441 Million cell updates/sec

Title: US-10-725-121A-103

Perfect score: 3090

Sequence: 1 MFGRKKRRVTSAPSNFHR.....LAKAGPPASIVPLMRQNR 591

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1096.5	35.5	540	T19956	hypothetical prote
2	1090.5	35.3	542	T19952	hypothetical prote
3	917	29.7	544	A57597	beta-p21-activated
4	907	29.4	544	I49376	p21 activated kina
5	895.5	29.0	525	S58682	protein kinase, p2
6	892.5	28.9	545	G01773	p21-activated, prot
7	892	28.9	544	S40482	serine/threonine-s
8	846.5	27.4	1230	T18259	probable serine/th
9	846.5	27.4	1230	T18259	serine/threonine p
10	837	27.1	939	S28394	probable serine/th
11	823.5	26.7	658	T39500	serine/threonine-s
12	814.5	26.4	658	S60170	protein kinase Pak
13	811.5	26.3	622	T15467	hypothetical prote
14	760	24.6	842	S60402	protein kinase CLA
15	722	23.4	378	T26684	hypothetical prote
16	697	22.6	589	T38086	serine/threonine-p
17	651.5	21.1	655	S51884	probable protein k
18	505	16.3	836	B96716	probable serine/th
19	498	16.1	471	T39232	probable serine th
20	478	15.5	653	JC6316	probable protein k
21	476	15.4	819	A53714	hypothetical prote
22	470.5	15.2	829	T29372	protein kinase (BC
23	464	15.0	436	S71886	hypothetical prote
24	463	15.0	1231	T18532	Ser20-like protein
25	462	15.0	1233	T14157	serine/threonine p
26	460	14.9	1206	T34021	protein kinase SK2
27	458.5	14.8	690	C96572	protein p12m16.4
28	455.5	14.7	1080	S48944	hypothetical prote

30	455	14.7	1233	2	T30989	serine/threonine p
31	450.5	14.6	312	2	T38525	serine/threonine p
32	445	14.4	1001	2	T17365	serine/threonine p
33	440.5	14.3	1075	2	T27623	hypothetical prote
34	440.5	14.3	1080	2	T27622	hypothetical prote
35	437.5	14.2	1062	2	S46367	protein kinase CDC
36	428	13.9	1228	2	T18897	hypothetical prote
37	424.5	13.7	659	1	A39723	protein kinase byr
38	421	13.6	553	2	T01479	hypothetical prote
39	419	13.6	883	2	A96662	hypothetical prote
40	417.5	13.5	1174	2	T43051	protein kinase C (
41	413	13.4	561	2	T51417	protein kinase-lik
42	409.5	13.3	652	2	T39722	serine/threonine p
43	409	13.2	891	2	T40503	protein kinase kin
44	405.5	13.1	1139	1	S61918	protein kinase C (
45	404.5	13.1	1135	1	A29813	132k ninaC protein

## ALIGNMENTS

RESULT 1									
T19956									
hypothetical protein C45B11.1b - Caenorhabditis elegans									
C/Species: Caenorhabditis elegans									
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 16-Aug-2004									
C/Accession: T19956									
R/McMurray, A.									
submitted to the EMBL Data Library, June 1996									
A/Reference number: Z19202									
A/Accession: T19956									
A/Status: preliminary; translated from GB/EMBL/DBJ									
A/Molecule type: DNA									
A/Residues: 1-540 <MIL>									
A/Cross-references: UNIPROT:Q9V3M1; EMBL:Z74029; P1DN:CAA98433.1; GSPDB:GN00023; CESP:C45									
A/Experimental source: clone C45B11									
C/Genetics:									
A/Gene: CESP:C45B11.1b									
A/Map position: 5									
A/Intons: 13/2; 62/3; 104/2; 231/3; 367/2; 468/3									
C/Superfamily: protein kinase homology									
Query Match									
Best Local Similarity 35.5%; Score 1096.5; DB 2; Length 540;									
Matches 251; Conservative 67; Mismatches 151; Indels 133; Gaps 9;									
Qy	4	KRKRRVTSAPSNFHRVHTGPDHOKFTGLPRNOSLI--ESARRPKPLVDPACTIS	61						
Db	29	RKYKSRISITPSNFERIHAGFDARSGTYTGLKQNALGPPRSISRPMVDPSCTTP	88						
Qy	62	IQGAPRTIVSGSKAGDGLTLTLDDEFNMSVTRNSLRLRDSPPPARARQRMDEBP	121						
Db	89	VDVAELKTVIRGSSSRNSPLPGFMNNSPMPSVARNSLRISAVASP-----	135						
Qy	122	ATTARAGPGKAGRGPRAGHSEAGGSGDRRRAPREKRPSSREGGPGQESSRDKRPLS	181						
Db	136	-----VVNVSARHSFRPLPLPVSGR-----	156						
Qy	182	GPVVTGPQAGLASGAKLAAGRPNTYPRADTDHPSKAGGGBHDVAVNGPSAGGLAIPQ	241						
Db	157	-----GTFP-----	170						
Qy	242	SSSSSRPPTRRAGADSPGYLGH-----ASBPQLAPPACTPAADAVPQPG	288						
Db	171	---RNQKPPW---STTFGVEKHQYQOITIVAPSRITTPQLQPKS-----PST	213						
Qy	289	PRSPQREP---QVSHSQPAALQVVDREDPSYLDNFKIGSGSTGYCIATVSSSK	345						
Db	214	PQMRQPPKCTEVSDBEFNALKFVVDGTPRSDLTLDYKQIGSGSTGVVEAAYKISTQ	273						
Qy	346	LVAIVKQDRLKQQRRLLEFENVVIMRDYQHENVEMNTSYLVGDELMVMEFLGGLATD	405						
Db	274	IVAVKRMNLKQQRRLLEFENVVIMRDYQHENVEMNTSYLVGDELMVMEFLGGLATD	333						

QY 406 IYHTRMNEBOIAVCLAVLQALSVLHAQGIYHRDIKSDSILLTHDGRVYKLSDFGFCQAV 465  
DB 334 IYATRTMTBEPQIATISRQVGLALDFLHARKVYIHRDIKSDSILLKRDGTVKLTDGFGCQL 393  
QY 466 SKEVPRKSLVGTPTWMAPELISRLPYGPEVDIWSLGIWVLEMDGEPFENEPPPLKAMK 525  
DB 394 SEEVPRRSLVGTPTWMAPELISRLPYGPEVDIWSLGIWVLEMDGEPFENEPPPLKAMK 453  
QY 526 MIRDNLPRRLKLNHKVSPSLKGLFDRLVYDPAQARATAELIKHPPFLAKAGPPASIVPLM 585  
DB 454 RIDDEHARFSRRAKVSVELSELISHCIYKDVNKRMPAKODLRHPPAKAQHSSSIAPLL 513  
QY 586 RQ 587  
DB 514 LQ 515

## RESULT 2

hypothetical protein C45B11.1a - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 16-Aug-2004  
C/Accession: T11952  
R/McMurray, A.  
Submitted to the EMBL Data Library, June 1996  
A/Reference number: Z19202  
A/Accession: T11952  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-542 <MIL>  
A/Cross-references: UNIPROT:Q16637, EMBL:Z74029, PIDN:CA098429.1, GSPDB:GN00023, CESP:C4  
A/Experimental source: clone C45B11  
A/Genetic8: CESP:C45B11.1a  
A/Map position: 5  
A/Intons: 13/2; 104/2; 233/3; 369/2; 470/3  
C/Superfamily: protein kinase homology

Query Match 35.3%; Score 1090.5; DB 2; Length 542;  
Best Local Similarity 42.0%; Pred. No. 6.6e-31;  
Matches 253; Conservative 69; Mismatches 149; Indels 131; Gaps 11;  
QY 4 KRKKRVISAPSNFEHRVHTGPDQHEQKFTGLPRQWOSLI--ESARRPKLVDPACITS 61  
DB 29 RKVKKSISTPSNFEHRIHAGFDARSCTYGLPKOMALLOPPRSISRPKMWDPSCITP 88  
QY 62 IQPGAPRTIVGSKGAKDGLTLLDDEFENNSVTRNSLRDSDPPPARAQENGMBEP 121  
DB 89 VDVAEKTVIRGP-----SSSFYNSPLP-----FGMINSPL 119  
QY 122 ATTARGPGKAGSGRPAHSEAGSGGDRRARAPBEKPKSREGSGGPOESSDKRPLS 181  
DB 120 MPSTV---AASNSLRISATSPVNVSSARHSFRPTLPVPSQR----- 158  
QY 182 GPDVGTQPPAGLAKLAAGRPFNTYPRADTDHPSRGAQGEPRHDVAENGPSAGLAIPO 241  
DB 159 -----GYFP-----NDSTYAPLPF-- 172  
QY 242 SSSSSRPPTRARGAPSPGVLPN-----ASEPOLAPPACTPAAPVPPPG 288  
DB 173 ---RNQRPW---STTFGEVKEPHQYQOITITVAPSRITTTQLQPKS-----PST 215  
QY 289 PRSQQRPE---QRVSHQFRAALQLVDPGDPRESYLDNFIKIGSGSTGIVCIATVRSAGK 345  
DB 216 PQARQOPKCTEGVSDEFNALKFVVDGTPRSDLTIDYKQIGSGSTGVVAAKYISTKQ 275  
QY 346 IYAVKQMDLRKQGRRELLFNEVIMRDYQHENNVEMVNSYLVGDELVMVMEFLREGALTD 405  
DB 276 IYAVKRMKLRKQGRRELLFNEVIMRDYQHENNVEMVNSYLVGDELVMVMEFLREGALTD 335  
QY 406 IYHTRMNEBOIAVCLAVLQALSVLHAQGIYHRDIKSDSILLTHDGRVYKLSDFGFCQAV 465

DB 336 IYATRTMTBEPQIATISRQVGLALDFLHARKVYIHRDIKSDSILLKRDGTVKLTDGFGCQL 393  
QY 466 SKEVPRKSLVGTPTWMAPELISRLPYGPEVDIWSLGIWVLEMDGEPFENEPPPLKAMK 525  
DB 396 SEEVPRRSLVGTPTWMAPELISRLPYGPEVDIWSLGIWVLEMDGEPFENEPPPLKAMK 453  
QY 526 MIRDNLPRRLKLNHKVSPSLKGLFDRLVYDPAQARATAELIKHPPFLAKAGPPASIVPLM 585  
DB 456 RIDDEHARFSRRAKVSVELSELISHCIYKDVNKRMPAKODLRHPPAKAQHSSSIAPLL 513  
QY 586 RQ 587  
DB 516 LQ 517

## RESULT 3

A57597  
beta-p21-activated protein kinase - rat  
N/Alternate names: beta-PAK  
C/Species: Rattus norvegicus (Norway rat)  
C/Date: 08-Feb-1996 #sequence\_revision 08-Feb-1996 #text\_change 09-Jul-2004  
C/Accession: A57597  
J./Manter, B./Chong, C./Zhao, Z.S./Leung, T./Michael, G./Hall, C./Lim, L.  
J. Biol. Chem. 270, 25070-25078, 1995  
A/Title: Molecular cloning of a new member of the p21-Cdc42/Rac-activated kinase (PAK) f  
A/Reference number: A57597; MUID:96027610; PMID:7559638  
A/Accession: A57597  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-544 <MAN>  
A/Cross-references: UNIPROT:Q62829; GB:U33314; NID:G1039424; PIDN:AAC52268.1; PID:G103942  
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo  
C/Keywords: ATP  
F/266-519/Domain: protein kinase homology <KIN>  
F/274-282/Region: protein kinase ATP-binding motif

Query Match 29.7%; Score 917; DB 2; Length 544;  
Best Local Similarity 35.8%; Pred. No. 5.2e-25;  
Matches 211; Conservative 87; Mismatches 165; Indels 126; Gaps 8;  
QY 4 KRKKRVISAPSNFEHRVHTGPDQHEQKFTGLPRQWOSLI--ESARRPKLVDP 56  
DB 63 KEKRPKISLPDSFEHRIHAGFDVAVGEPFGIEQWARRLQTSNITGLBKKNPQAVLD- 121  
QY 57 ACTISIQPGAPRTIVGSKGAKDGLTLLDDEFENNSVTRNSLRDSDPPPARAQENG 116  
DB 122 -----VKKFYDSKRYNNQRYMSFT----- 141  
QY 117 MPREPATTARGPGKAGSGRPAHSEAGSGGDRRARAPBEKPKSREGSGGPOESSRD 176  
DB 142 -----SGDSSAHGYIAAHQ-----SYTKTASBPPLAPPVSEEDDEBEDEED 184  
QY 177 KRPLSGPDVGTQPPAGLAKLAAGRPFNTYPRADTDHPSRGAQGEPRHDVAENGPSAGC 236  
DB 185 NEP---PPVIAAPREHTKS-----IYRSVVESIASPA-----APNKEATPP 223  
QY 237 LAIPQSSSSSRPPTRARGAPSPGVLPN-----ASEPOLAPPACTPAAPVPPPG 288  
DB 224 SAENANSTLYRNTDR-----ORKK 243  
QY 297 QRVSHEQFRAALQLVDPGDPRESYLDNFIKIGSGSTGIVCIATVRSAGKLVAVKQMDLRK 356  
DB 244 SKMTDEILKLSIVSGDPKKYTRFEKIKGGASCTVTYALDIAGQVAILKQNMLOQ 303  
QY 357 QQRRELLFNEVIMRDYQHENNVEMVNSYLVGDELVMVMEFLREGALTDIYHTRMNEQ 416  
DB 304 QPKKELIINEILVWREKNKNINIVYLDSTYLVGDELVMVMEFLREGALTDIYHTRMNEQ 363  
QY 417 IYAVCLAVLQALSVLHAQGIYHRDIKSDSILLTHDGRVYKLSDFGFCQAVKEVPRRSIV 476  
DB 364 IYAVCRECLQALPLFLSHNQVYHRDIKSDNILLKDGDSVVKLTDGFCQAVTPEQSKRTMV 423  
QY 477 GTPYMAPELISRLPYGPEVDIWSLGIWVLEMDGEPFENEPPPLKAMKIRDNLPRLK 536

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: November 17, 2005, 09:50:45 ; Search time 139.832 Seconds  
(without alignments)  
1634.641 Million cell updates/sec

Title: US-10-725-121A-103

Perfect score: 3090  
Sequence: 1 MFGRKKRKRKVISAPNSFRR.....LAKAGPPASIVPLMRQNRTR 591

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

1: geneseqp19808:\*  
2: geneseqp19908:\*  
3: geneseqp20008:\*  
4: geneseqp20018:\*  
5: geneseqp20028:\*  
6: geneseqp20038:\*  
7: geneseqp20038:\*  
8: geneseqp20048:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARYS

Result No.	Score	Query Match	Length	ID	Description
1	3090	100.0	591	2	AAVS5964 Pull leng
2	3090	100.0	591	3	AAVS5128 Human ser
3	3090	100.0	591	7	ADCS3707 Nuclear f
4	3090	100.0	591	7	ADDB8973 Human can
5	3090	100.0	591	7	ADFA5079 Human kin
6	3090	100.0	591	8	ADH23362 Human p21
7	3090	100.0	591	8	ABM82101 Human p21
8	3090	100.0	620	8	ABG19308 Novel hum
9	3075	99.5	588	8	ADH19308 Novel hum
10	2860	92.6	593	8	ADJ96654 Human p21
11	2674.5	86.6	522	8	ABM84642 Human ste
12	2674.5	86.6	522	8	ABM84644 Human dia
13	2674.5	86.6	522	8	ABM84644 Human dia
14	2662	86.1	517	8	ABM84648 Human dia
15	2662	86.1	517	8	ABM84647 Human dia
16	2662	86.1	517	8	ABM84643 Human dia
17	2662	86.1	517	8	ABM84646 Human dia
18	2558	82.8	501	7	ADCS3709 Nuclear f
19	2302.5	71.3	438	8	ADCS3705 Nuclear f
20	2194.5	71.0	438	8	ADCS3763 Human kin
21	2073	67.1	398	2	AAVS5941 Human PAK
22	1748	56.6	719	7	ADJ83007 Murine ma
23	1739	56.3	719	4	AAVS8963 Human pol
24	1734	56.1	719	4	AAVS8963 Human pol
25	1731	56.0	719	4	AAVS5705 Novel pro

26	1731	56.0	719	7	ADCS37451 Nuclear f
27	1731	56.0	719	7	ADFA5080 Human kin
28	1731	56.0	719	8	ADJ29314 Human MAR
29	1730	56.0	719	8	AAE02187 Human p21
30	1549	50.1	457	8	ADH42207 Novel hum
31	1549	50.1	457	8	ADH42215 Novel hum
32	1549	50.1	457	8	ADH42217 Novel hum
33	1545	50.0	457	8	ADH42213 Novel hum
34	1535	49.7	457	8	ADH42211 Novel hum
35	1502	48.6	293	7	ADBE15851 Human pro
36	1497.5	48.5	639	8	ADH23364 Fruit fly
37	1494.5	48.4	635	8	ADH23360 Fruit fly
38	1473.5	47.7	681	2	AAVS5940 Human PAK
39	1473.5	47.7	681	2	AAVS5967 Signal tr
40	1473.5	47.7	681	4	AAE20337 Human PAK
41	1473.5	47.7	681	5	ABP64709 Human pro
42	1473.5	47.7	681	8	ADQ15150 Human can
43	1473.5	47.7	701	6	ABU11508 Human MDD
44	1470.5	47.6	681	3	AAVS03970 Mutant si
45	1470.5	47.6	681	4	AAVS63230 Amino aci

## ALIGNMENTS

RESULT 1	AAVS5964 standard; protein; 591 AA.
AAVS5964	
XX	
AC	AAVS5964;
XX	
DT	18-FEB-2000 (first entry)
XX	
DE	Full length human PAKS protein.
XX	
KW	Antitumorigenic; antiproliferative; antiinflammatory; antiallergic; osteopathic;
KW	antiproliferative; antiproliferative; antiproliferative; antiproliferative;
KW	neuroprotective; cardioprotective; cytoprotective; antidiabetic;
KW	vulnerability; STB20; protein kinase; STK2; STK3; STK4; STK5; STK6; STK7;
KW	ZC1; ZC2; ZC3; ZC4; KHS2; STU1; STU2; STU3; STU4; STU5; STU6; STU7;
KW	antibody; gene therapy; rheumatoid arthritis; arteriosclerosis; asthma;
KW	inflammatory bowel disease; Crohn's disease; osteoarthritis; psoriasis;
KW	rhinitis; autoimmunity; organ transplantation; multiple sclerosis;
KW	myocardial infarction; cardiovascular disease; stroke; renal failure;
KW	oxidative stress-related neurodegenerative disorder; Parkinson's disease;
KW	amyotrophic lateral sclerosis; Leigh syndrome; cancer; cardiomyopathy;
KW	ischemic disorder; inflammation; diabetes mellitus; fibrosis; microsis;
KW	mesangial disorder; growth regulation; wound healing; T cell activation;
KW	immunopressant.
XX	
OS	Homo sapiens.
XX	
PN	WO9553036-A2.
XX	
PD	21-OCT-1999.
XX	
PF	13-APR-1999; 99WO-US008150.
XX	
PR	14-APR-1998; 98US-0081784P.
XX	
PA	(SUGR-) SUGEN INC.
XX	
PI	Plozman G, Martinez R, Whyte D;
XX	
DR	WPI; 1999-611301/52.
XX	
PT	N-PDB; AA240538.
XX	
PS	Novel kinase-related polypeptides used for the diagnosis and treatment of
XX	kinase-related diseases and disorders.
XX	Claim 11, Page 366-368; 387pp; English.
CC	This sequence represents a novel STB20-related protein kinase. The

invention relates to nucleic acid molecule encoding a kinase polypeptide selected from SULK2, SULK3, SULK4, SULK5, SULK6, SULK7, ZC1, ZC2, ZC3, ZC4, KHS2, SUVU1, SUVU3, GSK2, PAK4 and PAK5. The proteins are used to identify agonists and antagonists, and to raise antibodies. The polynucleotides are useful in gene therapy protocols. The polynucleotides, antibodies, antagonists and agonists may be used to treat diseases such as immune-related disorders and diseases (e.g. rheumatoid arthritis, atherosclerosis, chronic inflammatory bowel disease (e.g. Crohn's disease), asthma, osteoarthritis, psoriasis, atherosclerosis, rhinitis, autoimmunity, and organ transplantation, chronic inflammatory disease, multiple sclerosis, organ transplantation, myocardial infarction, cardiovascular disease, stroke, renal failure, oxidative stress-related neurodegenerative disorders (e.g. amyotrophic lateral sclerosis, Parkinson's disease and Leigh syndrome), cancer, cardiomyopathies, ischemic disorders, inflammatory disorders, diabetes mellitus, fibrotic and meangial disorders. The proteins may also be useful for cell growth regulation (e.g. in wound healing), T cell activation, mitosis control, and as immunosuppressants

Sequence 591 AA:

Query Match 100.0%; Score 3090; DB 2; Length 591;

Best Local Similarity 100.0%; Pred. No. 2.8e-216; Mismatches 0; Gaps 0;

Matches 591; Conservative 0; Indels 0; Gaps 0;

1 MFGKRRKRVISAPNFEHVRHTGPDHOKFTGLPRQMSLIESARRPKPLVDPACIT 60  
1 MFGKRRKRVISAPNFEHVRHTGPDHOKFTGLPRQMSLIESARRPKPLVDPACIT 60  
61 SIQGAPKTIYRSGKAGKADGALLTLDEFENMSYTRNSLRSDPPPPARARQNGMPEE 120  
61 SIQGAPKTIYRSGKAGKADGALLTLDEFENMSYTRNSLRSDPPPPARARQNGMPEE 120  
121 PATTARCGPGKAGRGPRFAGHSEAGSGSGDRRRAPKRRPKSSREGSGGPOESSRDKRPL 180  
121 PATTARCGPGKAGRGPRFAGHSEAGSGSGDRRRAPKRRPKSSREGSGGPOESSRDKRPL 180  
121 PATTARCGPGKAGRGPRFAGHSEAGSGSGDRRRAPKRRPKSSREGSGGPOESSRDKRPL 180  
181 SGPVGTGPPAGLASGAKLAAGRPNTYPRADTHPSRGAQGEHPDVAHPNGSAGLAIP 240  
181 SGPVGTGPPAGLASGAKLAAGRPNTYPRADTHPSRGAQGEHPDVAHPNGSAGLAIP 240  
181 SGPVGTGPPAGLASGAKLAAGRPNTYPRADTHPSRGAQGEHPDVAHPNGSAGLAIP 240  
241 QSSSSSRPPTARAGAPSPGVLGPHASEPQLAPACTPAAVAVGPPGPPSPQREPORVS 300  
241 QSSSSSRPPTARAGAPSPGVLGPHASEPQLAPACTPAAVAVGPPGPPSPQREPORVS 300  
301 HEQFRALQLVVDGDRSYLDNFIKIGESTGIVCIATYRSSGKLVAVKMDLRKQORR 360  
301 HEQFRALQLVVDGDRSYLDNFIKIGESTGIVCIATYRSSGKLVAVKMDLRKQORR 360  
301 HEQFRALQLVVDGDRSYLDNFIKIGESTGIVCIATYRSSGKLVAVKMDLRKQORR 360  
361 ELTFNEVIMRDYGHENVMYNSYLVDELAVMVEFLGEGALTDIYTHMNEBQIAAV 420  
361 ELTFNEVIMRDYGHENVMYNSYLVDELAVMVEFLGEGALTDIYTHMNEBQIAAV 420  
361 ELTFNEVIMRDYGHENVMYNSYLVDELAVMVEFLGEGALTDIYTHMNEBQIAAV 420  
421 CLAVIALSVLAAGVTHRDIKSDSILLTHDGRVGLSDPGCAQVSKVPPRKSILVETPY 480  
421 CLAVIALSVLAAGVTHRDIKSDSILLTHDGRVGLSDPGCAQVSKVPPRKSILVETPY 480  
421 CLAVIALSVLAAGVTHRDIKSDSILLTHDGRVGLSDPGCAQVSKVPPRKSILVETPY 480  
481 WMAEELISRLPYGEBVDIWSIGIWIEMVDEPPYFNRPPLKAMKMTLDNLPRLKULHK 540  
481 WMAEELISRLPYGEBVDIWSIGIWIEMVDEPPYFNRPPLKAMKMTLDNLPRLKULHK 540  
541 VSPSLKGFIDRLVLRDPAQRATTAELKHPFLAKAGPASIIVPLMKQKTR 591  
541 VSPSLKGFIDRLVLRDPAQRATTAELKHPFLAKAGPASIIVPLMKQKTR 591

RESULT 2  
AAV59128  
ID AAV59128 standard; protein; 591 AA.XX AAV59128;  
AC  
XX 08-MAR-2000 (first entry)

XX Human serine/threonine kinase, PAK4.  
DB PAK4, serine/threonine kinase; GTPase; intracellular signal cascade; Rac;  
XX Cdc42H; morphogenesis; mitogenesis; JNK; p38 MAP kinase; human;  
XX actin polymerization; filopodia; cancer; arthritis.  
OS Homo sapiens.  
XX W09963073-A1.  
XX 09-DEC-1999.  
XX 21-MAY-1999; 99WO-US011341.  
XX 21-MAY-1998; 98US-00082737.  
XX (UYCO ) UNIV COLUMBIA NEW YORK.  
XX Minden A;  
XX WPI; 2000-072881/06.  
XX N-PSDB; AAZ40657.  
XX Novel mammalian nucleic acid useful for treating cancer and arthritis.  
XX Claim 8; Fig 1A-B; 95pp; English.

The invention relates to an isolated mammalian nucleic acid that encodes PAK4, a novel serine/threonine kinase or its mutant homolog. PAK4 is an effector for the GTPases Rac and Cdc42Hs which are involved in intracellular signal cascades, morphogenesis and mitogenesis, and activate the JNK and p38 MAP kinase pathways. Inhibiting interaction of PAK4 with these enzymes will thus result in inhibition of actin polymerization and formation of filopodia. The PAK4 nucleic acid used for recombinant production of the protein, and as a source of probes for identifying homologous sequences and of (anti)sense oligonucleotides for inhibiting PAK4 expression. The protein, or its fragments, are used to raise specific antibodies and these are useful as ligands for therapeutic inhibition of interaction between PAK4 and its native binding partners. CC Inhibition of PAK4 activity or expression is used for treatment of cancer and arthritis. The present sequence represents the human serine/threonine kinase, PAK4

Sequence 591 AA:

Query Match 100.0%; Score 3090; DB 3; Length 591;

Best Local Similarity 100.0%; Pred. No. 2.8e-216; Mismatches 0; Gaps 0;

Matches 591; Conservative 0; Indels 0; Gaps 0;

1 MFGKRRKRVISAPNFEHVRHTGPDHOKFTGLPRQMSLIESARRPKPLVDPACIT 60  
1 MFGKRRKRVISAPNFEHVRHTGPDHOKFTGLPRQMSLIESARRPKPLVDPACIT 60  
61 SIQGAPKTIYRSGKAGKADGALLTLDEFENMSYTRNSLRSDPPPPARARQNGMPEE 120  
61 SIQGAPKTIYRSGKAGKADGALLTLDEFENMSYTRNSLRSDPPPPARARQNGMPEE 120  
121 PATTARCGPGKAGRGPRFAGHSEAGSGSGDRRRAPKRRPKSSREGSGGPOESSRDKRPL 180  
121 PATTARCGPGKAGRGPRFAGHSEAGSGSGDRRRAPKRRPKSSREGSGGPOESSRDKRPL 180  
121 PATTARCGPGKAGRGPRFAGHSEAGSGSGDRRRAPKRRPKSSREGSGGPOESSRDKRPL 180  
181 SGPVGTGPPAGLASGAKLAAGRPNTYPRADTHPSRGAQGEHPDVAHPNGSAGLAIP 240  
181 SGPVGTGPPAGLASGAKLAAGRPNTYPRADTHPSRGAQGEHPDVAHPNGSAGLAIP 240  
181 SGPVGTGPPAGLASGAKLAAGRPNTYPRADTHPSRGAQGEHPDVAHPNGSAGLAIP 240  
241 QSSSSSRPPTARAGAPSPGVLGPHASEPQLAPACTPAAVAVGPPGPPSPQREPORVS 300  
241 QSSSSSRPPTARAGAPSPGVLGPHASEPQLAPACTPAAVAVGPPGPPSPQREPORVS 300  
301 HEQFRALQLVVDGDRSYLDNFIKIGESTGIVCIATYRSSGKLVAVKMDLRKQORR 360  
301 HEQFRALQLVVDGDRSYLDNFIKIGESTGIVCIATYRSSGKLVAVKMDLRKQORR 360



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2005, 09:51:36 ; Search time 39.4398 Seconds  
(without alignments)  
1118.605 Million cell updates/sec

Title: US-10-725-121A-103

Perfect score: 3090  
Sequence: 1 MFGRKKRVRISAPSNPEHR.....LAKAGPPASTVPLMRQNRTR 591

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74643064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3090	100.0	591	3 US-09-082-737-2	Sequence 2, Appl1
2	3090	100.0	591	4 US-09-688-1888-103	Sequence 103, App
3	3090	100.0	591	4 US-09-718-032-2	Sequence 2, Appl1
4	3090	100.0	591	4 US-09-291-417D-103	Sequence 103, App
5	3090	100.0	591	4 US-09-949-016-6665	Sequence 6665, Ap
6	3090	100.0	620	4 US-09-949-016-7206	Sequence 7206, Ap
7	2073	67.1	398	4 US-09-688-1888-30	Sequence 30, Appl1
8	2073	67.1	398	4 US-09-291-417D-30	Sequence 30, Appl1
9	1473.5	47.7	681	4 US-09-688-1888-29	Sequence 29, Appl1
10	1473.5	47.7	681	4 US-09-291-417D-29	Sequence 29, Appl1
11	1473.5	47.7	684	4 US-09-949-016-10445	Sequence 10445, A
12	1470.5	47.6	681	4 US-09-765-815-2	Sequence 2, Appl1
13	1277	41.3	250	3 US-09-082-737-3	Sequence 3, Appl1
14	1277	41.3	250	4 US-09-718-032-2	Sequence 2, Appl1
15	912	29.5	544	2 US-08-935-760-2	Sequence 2, Appl1
16	909	29.4	544	4 US-09-949-016-11562	Sequence 11562, A
17	907	29.4	544	4 US-09-688-1888-95	Sequence 95, Appl1
18	907	29.4	544	4 US-09-291-417D-95	Sequence 95, Appl1
19	904	29.1	524	4 US-09-538-092-1301	Sequence 1301, Ap
20	900	29.1	524	2 US-08-615-942A-2	Sequence 2, Appl1
21	900	29.1	524	4 US-09-237-325-2	Sequence 2, Appl1
22	899	29.1	544	3 US-08-559-397A-19	Sequence 19, Appl1
23	894	28.9	506	1 US-08-359-780-2	Sequence 2, Appl1
24	894	28.9	506	1 US-08-475-682-2	Sequence 2, Appl1
25	894	28.9	506	1 US-08-780-833-2	Sequence 2, Appl1
26	894	28.9	506	1 US-08-636-036-2	Sequence 2, Appl1
27	894	28.9	506	3 US-08-918-509-2	Sequence 2, Appl1

28	894	28.9	506	3 US-09-108-262-2	Sequence 2, Appl1
29	894	28.9	506	4 US-09-688-1888-94	Sequence 94, Appl1
30	894	28.9	506	4 US-09-291-417D-94	Sequence 94, Appl1
31	893.5	28.9	551	4 US-09-949-016-10951	Sequence 10951, A
32	892.5	28.9	545	4 US-09-538-092-1297	Sequence 1297, Ap
33	891.5	28.9	545	2 US-08-935-760-4	Sequence 4, Appl1
34	891.5	28.9	545	4 US-09-688-1888-93	Sequence 93, Appl1
35	891.5	28.9	544	4 US-09-291-417D-93	Sequence 93, Appl1
36	867	28.1	544	3 US-08-559-397A-29	Sequence 29, Appl1
37	849	27.5	544	3 US-08-559-397A-30	Sequence 30, Appl1
38	833.5	27.0	465	2 US-08-114-555A-2	Sequence 2, Appl1
39	821	26.6	694	3 US-08-559-397A-31	Sequence 31, Appl1
40	814.5	26.4	465	3 US-08-559-397A-2	Sequence 2, Appl1
41	804	26.0	305	4 US-09-765-815-10	Sequence 10, Appl1
42	782	25.3	793	3 US-09-588-256-10	Sequence 10, Appl1
43	769	24.9	268	2 US-08-852-743-3	Sequence 3, Appl1
44	769	24.9	268	3 US-09-185-370-3	Sequence 3, Appl1
45	755	24.4	410	4 US-09-248-796A-18461	Sequence 18461, A

## ALIGNMENTS

```
RESULT 1
US-09-082-737-2
Sequence 2, Application US/09082737
Patent No. 6013500
GENERAL INFORMATION:
APPLICANT: Minden, Audrey
TITLE OF INVENTION: PAK4, A No. 6013500e1 Gene Encoding A Serine/
TITLE OF INVENTION: Threonine Kinase
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES:
ADDRESSER: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 11230
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/082,737
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/55311
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 591 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-082-737-2
Query Match 100.0%; Score 3090; DB 3; Length 591;
Best Local Similarity 100.0%; Pred. No. 36-191;
Matches 591; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MFGRKKRVRISAPSNPEHRVHTGPOHEKFTGLPRQOSLIBESARRKXPLVDPACIT 60
DB 1 MFGRKKRVRISAPSNPEHRVHTGPOHEKFTGLPRQOSLIBESARRKXPLVDPACIT 60
QY 61 SIQGAPKTVRSKAKDGLTLLDPEFNMVTSNLSRRDSPPPARARQENGPEE 120
```

Db 61 SIOPGAPKTIYVRSKGAKDGLATLLDDEFENMSTVTRNSLRRDSPPPARARQENGMPEE 120  
Qy 121 PATTARGPAGKASRGGRFAGHSEAGGSGDRRRAGPEKRPKSSRESGGSGQESSRDRPL 180  
Db 121 PATTARGPAGKASRGGRFAGHSEAGGSGDRRRAGPEKRPKSSRESGGSGQESSRDRPL 180  
Qy 181 SGPVGTGPAGLASGAKLAAGRPFNTYPRADTDHPSRGAQGEBHDAVAPNGPSAGLAIP 240  
Db 181 SGPVGTGPAGLASGAKLAAGRPFNTYPRADTDHPSRGAQGEBHDAVAPNGPSAGLAIP 240  
Qy 241 QSSSSSSRPPTRRAGAPSPVLPAGHASEPQLAPACTPAAVAVGPPGPSPOREPORVS 300  
Db 241 QSSSSSSRPPTRRAGAPSPVLPAGHASEPQLAPACTPAAVAVGPPGPSPOREPORVS 300  
Qy 301 HEOPRALQVLPDGPDRSYLDNFIKIGESTGIVCIATVRSRSGKLVAVKMDLRKQORR 360  
Db 301 HEOPRALQVLPDGPDRSYLDNFIKIGESTGIVCIATVRSRSGKLVAVKMDLRKQORR 360  
Qy 361 ELLEFNEVIMRDYOHENNVEMVNSYLVGDELVMVMEFLGEGALTDIYTHTRMNEQJAAV 420  
Db 361 ELLEFNEVIMRDYOHENNVEMVNSYLVGDELVMVMEFLGEGALTDIYTHTRMNEQJAAV 420  
Qy 421 CLAVLQALSVLHAQGVYIHRDIKSDSILLTHDGRVKLSDFGCAQVSKVPRRKSIVGTPT 480  
Db 421 CLAVLQALSVLHAQGVYIHRDIKSDSILLTHDGRVKLSDFGCAQVSKVPRRKSIVGTPT 480  
Qy 481 WMAPELISRLPYGGEVDIWSLGIWVEMVGEPPYFNEPPLKAMKMRDNLPRLLKXILHK 540  
Db 481 WMAPELISRLPYGGEVDIWSLGIWVEMVGEPPYFNEPPLKAMKMRDNLPRLLKXILHK 540  
Qy 541 VSPSLKGFLLRLVLRDPAQRATAELKHPFLAKAGPASTIVPLMRQNRTR 591  
Db 541 VSPSLKGFLLRLVLRDPAQRATAELKHPFLAKAGPASTIVPLMRQNRTR 591

RESULT 2  
US-09-688-103  
; Sequence 103, Application US/09688188B  
; Patent No. 6656716  
; GENERAL INFORMATION:  
; APPLICANT: PLOMMAN, GREGORY  
; APPLICANT: MARTINEZ, RICARDO  
; APPLICANT: WHITE, DAVID  
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES  
; PILE REFERENCE: 038602/0328  
; CURRENT APPLICATION NUMBER: US/09/688, 1888  
; CURRENT FILING DATE: 2000-10-16  
; PRIOR APPLICATION NUMBER: 09/291,417  
; PRIOR FILING DATE: 1999-04-14  
; PRIOR APPLICATION NUMBER: 60/081,784  
; PRIOR FILING DATE: 1998-04-14  
; NUMBER OF SEQ ID NOS: 155  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 103  
; LENGTH: 591  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-688-188B-103

Query Match 100.0%; Score 3090; DB 4; Length 591;  
Best Local Similarity 100.0%; Pred. No. 3e-191;  
Matches 591; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MFGRRKKRVEISAPSNFEHRVHTGFDQHEOKFTGLPRQWOSLIBESARRPKPLVDPACT 60  
Db 1 MFGRRKKRVEISAPSNFEHRVHTGFDQHEOKFTGLPRQWOSLIBESARRPKPLVDPACT 60  
Qy 61 SIOPGAPKTIYVRSKGAKDGLATLLDDEFENMSTVTRNSLRRDSPPPARARQENGMPEE 120  
Db 61 SIOPGAPKTIYVRSKGAKDGLATLLDDEFENMSTVTRNSLRRDSPPPARARQENGMPEE 120  
Qy 121 PATTARGPAGKASRGGRFAGHSEAGGSGDRRRAGPEKRPKSSRESGGSGQESSRDRPL 180  
Db 121 PATTARGPAGKASRGGRFAGHSEAGGSGDRRRAGPEKRPKSSRESGGSGQESSRDRPL 180

Db 121 PATTARGPAGKASRGGRFAGHSEAGGSGDRRRAGPEKRPKSSRESGGSGQESSRDRPL 180  
Qy 181 SGPVGTGPAGLASGAKLAAGRPFNTYPRADTDHPSRGAQGEBHDAVAPNGPSAGLAIP 240  
Db 181 SGPVGTGPAGLASGAKLAAGRPFNTYPRADTDHPSRGAQGEBHDAVAPNGPSAGLAIP 240  
Qy 241 QSSSSSSRPPTRRAGAPSPVLPAGHASEPQLAPACTPAAVAVGPPGPSPOREPORVS 300  
Db 241 QSSSSSSRPPTRRAGAPSPVLPAGHASEPQLAPACTPAAVAVGPPGPSPOREPORVS 300  
Qy 301 HEOPRALQVLPDGPDRSYLDNFIKIGESTGIVCIATVRSRSGKLVAVKMDLRKQORR 360  
Db 301 HEOPRALQVLPDGPDRSYLDNFIKIGESTGIVCIATVRSRSGKLVAVKMDLRKQORR 360  
Qy 361 ELLEFNEVIMRDYOHENNVEMVNSYLVGDELVMVMEFLGEGALTDIYTHTRMNEQJAAV 420  
Db 361 ELLEFNEVIMRDYOHENNVEMVNSYLVGDELVMVMEFLGEGALTDIYTHTRMNEQJAAV 420  
Qy 421 CLAVLQALSVLHAQGVYIHRDIKSDSILLTHDGRVKLSDFGCAQVSKVPRRKSIVGTPT 480  
Db 421 CLAVLQALSVLHAQGVYIHRDIKSDSILLTHDGRVKLSDFGCAQVSKVPRRKSIVGTPT 480  
Qy 481 WMAPELISRLPYGGEVDIWSLGIWVEMVGEPPYFNEPPLKAMKMRDNLPRLLKXILHK 540  
Db 481 WMAPELISRLPYGGEVDIWSLGIWVEMVGEPPYFNEPPLKAMKMRDNLPRLLKXILHK 540  
Qy 541 VSPSLKGFLLRLVLRDPAQRATAELKHPFLAKAGPASTIVPLMRQNRTR 591  
Db 541 VSPSLKGFLLRLVLRDPAQRATAELKHPFLAKAGPASTIVPLMRQNRTR 591

RESULT 3  
US-09-718-032-2  
; Sequence 2, Application US/09718032  
; Patent No. 667168  
; GENERAL INFORMATION:  
; APPLICANT: The Trustees of Columbia University  
; APPLICANT: Minden, Audrey  
; TITLE OF INVENTION: PAK4, A NOVEL GENE ENCODING A SERINE/THREONINE KINASE  
; PILE REFERENCE: 575/55311-A-PCT-US  
; CURRENT APPLICATION NUMBER: US/09/718, 032  
; CURRENT FILING DATE: 2000-11-21  
; PRIOR APPLICATION NUMBER: PCT/US99/11341  
; PRIOR FILING DATE: 1999-05-21  
; PRIOR APPLICATION NUMBER: 09/082,737  
; PRIOR FILING DATE: 1998-05-21  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 591  
; TYPE: PRT  
; ORGANISM: human  
US-09-718-032-2

Query Match 100.0%; Score 3090; DB 4; Length 591;  
Best Local Similarity 100.0%; Pred. No. 3e-191;  
Matches 591; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MFGRRKKRVEISAPSNFEHRVHTGFDQHEOKFTGLPRQWOSLIBESARRPKPLVDPACT 60  
Db 1 MFGRRKKRVEISAPSNFEHRVHTGFDQHEOKFTGLPRQWOSLIBESARRPKPLVDPACT 60  
Qy 61 SIOPGAPKTIYVRSKGAKDGLATLLDDEFENMSTVTRNSLRRDSPPPARARQENGMPEE 120  
Db 61 SIOPGAPKTIYVRSKGAKDGLATLLDDEFENMSTVTRNSLRRDSPPPARARQENGMPEE 120  
Qy 121 PATTARGPAGKASRGGRFAGHSEAGGSGDRRRAGPEKRPKSSRESGGSGQESSRDRPL 180  
Db 121 PATTARGPAGKASRGGRFAGHSEAGGSGDRRRAGPEKRPKSSRESGGSGQESSRDRPL 180  
Qy 181 SGPVGTGPAGLASGAKLAAGRPFNTYPRADTDHPSRGAQGEBHDAVAPNGPSAGLAIP 240  
Db 181 SGPVGTGPAGLASGAKLAAGRPFNTYPRADTDHPSRGAQGEBHDAVAPNGPSAGLAIP 240